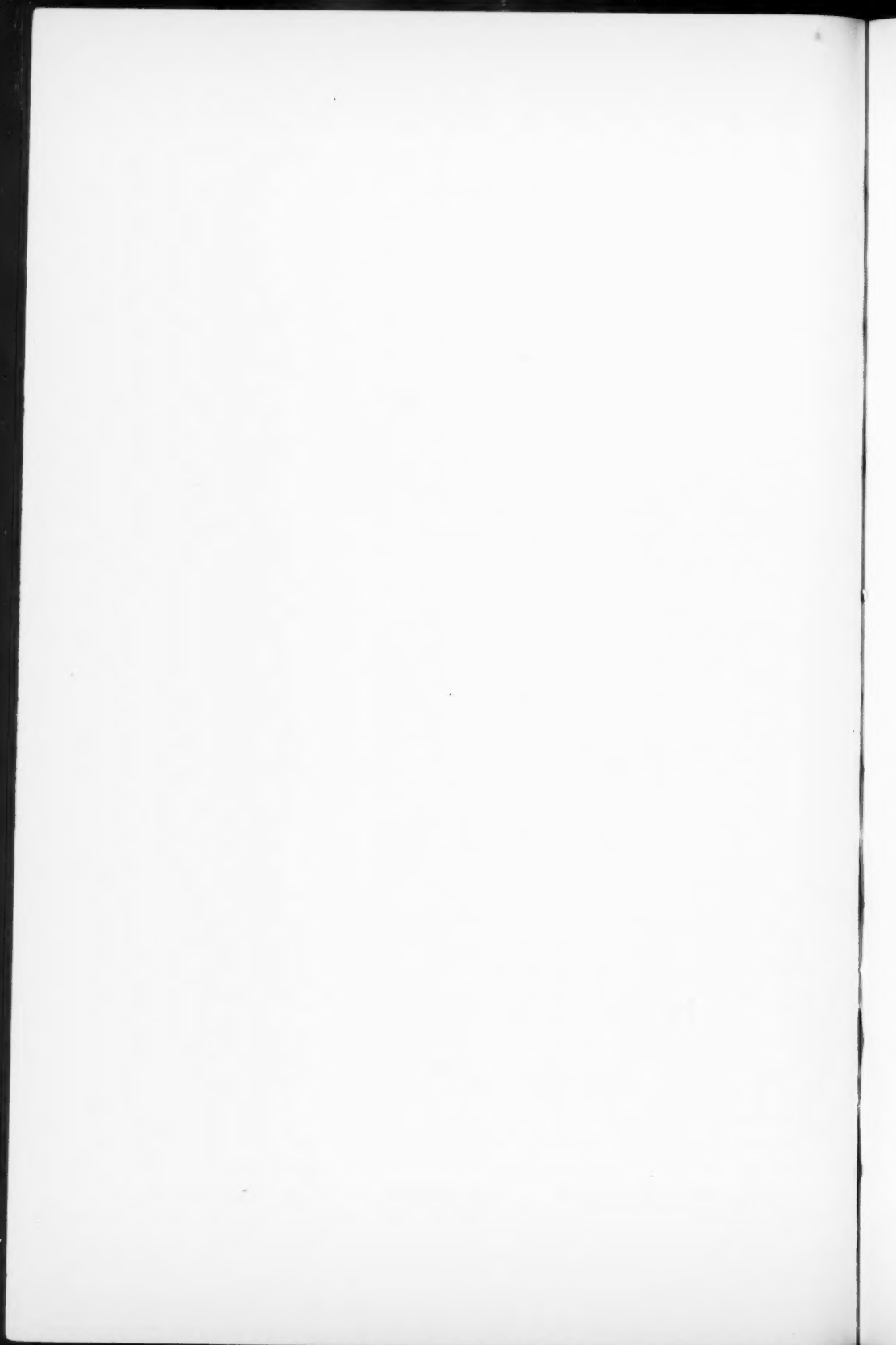


# Psychometrika

## CONTENTS

A TWO-SAMPLE TEST . . . . .	239
LINCOLN E. MOSES	
A NOTE ON CLUSTER-DIRECTED ANALYSIS . . . . .	249
C. J. ADCOCK	
METHODS OF SOLVING SOME PERSONNEL-CLASSIFICATION PROBLEMS . . . . .	255
D. F. VOTAW, JR.	
THE CLOSURE FACTORS RELATED TO OTHER COGNITIVE PROCESSES . . . . .	267
CAROL PEMBERTON	
SOME EXAMPLES OF MULTIVARIATE ANALYSIS IN EDUCATIONAL AND PSYCHOLOGICAL RESEARCH . . . .	289
FRED T. TYLER	
NOTES ON A PROBLEM OF MULTIPLE CLASSIFICATION .	297
FREDERIC M. LORD	
THE ROLLING TOTALS METHOD OF COMPUTING SUMS, SUMS OF SQUARES, AND SUMS OF CROSS-PRODUCTS	305
J. DOUGLAS AYERS AND J. PERHAM STANLEY	
AN EXTENDED TABLE OF CHI-SQUARE FOR TWO DEGREES OF FREEDOM, FOR USE IN COMBINING PROBABILITIES FROM INDEPENDENT SAMPLES . . . . .	311
MORDECAI H. GORDON, EDWARD H. LOVELAND, AND EDWARD E. CURETON	
ORTHOGONAL AND OBLIQUE SIMPLE STRUCTURE . . . .	317
W. A. GIBSON	
COMPANION NOMOGRAPHS FOR TESTING THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN UNCORRELATED PERCENTAGES . . . . .	325
VALENTINE APPEL	
ON THE ECONOMIES OF A PRE-SCREENING TECHNIQUE FOR APTITUDE TEST BATTERIES . . . . .	331
A. G. ARBOUS AND H. S. SICHEL	
THE FACTORIAL COMPOSITION AND VALIDITY OF DIFFERENTLY SPEEDED TESTS . . . . .	347
CHARLES T. MYERS	
PAUL S. DWYER, <i>Linear Computations</i> . . . . .	353
A Review by HARRY H. HARMAN	
BOOKS RECEIVED . . . . .	355
PSYCHOMETRIC MONOGRAPHS ANNOUNCEMENT . . . .	357



## A TWO-SAMPLE TEST

LINCOLN E. MOSES

TEACHERS COLLEGE, COLUMBIA UNIVERSITY

In certain investigations involving an experimental group and a control group, the effect of the experimental treatment may be expected to manifest itself by increasing the scores of some subjects but by decreasing the scores of other subjects. Examples arise in the area of defensive behavior. Customary tests of differences between means are inappropriate for assessing the existence of effect of the experimental treatment in such cases. A test based on the ranks of the observations is proposed; it will be sensitive to extreme scores (either large or small or both) for experimentals.

Often in psychological studies where behavior can be represented by scores along a continuum, the import of extremely large or extremely small scores is much the same. For instance, in certain association or recognition situations either a rapid or a delayed response is evidence of defensive behavior. Similar considerations apply for many "personality" variables. For convenience of exposition we shall employ terminology special to perceptual defense.

Suppose that an investigator wishes to determine whether the behavior in one group (experimentals) is defensive, as contrasted with the behavior of another group (controls). He then faces the difficulty that statistical techniques based upon differences of means (or medians, or mean ranks, etc.) may be inherently invalid for the problem. Suppose that on each subject is obtained one measurement which, if small, indicates a defensive response sometimes characterized as "vigilant." If the number is large it represents a defensive response of a character sometimes called "repressive." The normal response is expected not to be extreme in either direction. Now if the experimental group consists of subjects, some of whom give "vigilant" defensive responses and some "repressive" defensive responses, their mean response may be quite close to the mean response of the controls. The null hypothesis might well be accepted when it should be rejected.

Since a test of means is not in order, the experimenter might propose to use the Wald-Wolfowitz run test. He would not be much better off than he was before, since the run test will merely tell him whether or not the two sets of measurements can be viewed as coming from a common population. If the result is that they can not, he still has no firm basis for stating that the experimentals exhibited defense as compared with the controls. For instance, suppose that all scores are arranged in order of increasing size and that one-

half of the control group scores lay at the beginning and were followed by all the experimental group scores which in turn were followed by the remaining control group scores. The run test here would reject the hypothesis of a common population but the character of the sample would certainly not justify the statement that the experimentals had given defensive responses; rather it would indicate that the controls had.

It would seem, then, that a suitable test is needed. Proposed in this paper is a test of the null hypothesis that the two groups, experimental and control, come from a common population, against the alternative that the experimentals give defensive responses. (More generally the alternative hypothesis is that the experimentals are "extreme" in one or both directions, relative to the controls.) The test is completely distribution-free. It will have its specified significance level whatever be the form of the common population, when the null hypothesis is true. It is not argued that this is the best test of the hypothesis which can be given, in the sense that presumably more sensitive ones could be constructed. The test has the desirable property that there is not an implicit assumption of a measurement scale with equal intervals.

It should be emphasized that the test to be proposed is specifically designed for the case where the best information of the investigator indicates that the experimental condition may affect the scores of some (or possibly all) subjects in one way and the scores of some (or possibly all) subjects in the opposite way. In this case the customary tests of location (or "slippage") based on means or medians or rank sums cannot be relied upon to give reasonable answers even with very large samples. However, if there exist *a priori* grounds for expecting that the effect (if any) of the experimental condition must be in the same direction for all subjects affected, then a suitable test of location is not only appropriate, but is to be preferred to the test here proposed.

For definiteness, let there be  $m$  observations from the control group which we will call  $x$ 's and let there be  $n$  observations from the experimental group which we will call  $y$ 's. Let the  $m + n$  observations be arranged in order of increasing size. If the null hypothesis that the  $x$ 's and  $y$ 's come from a common population is true, we should expect the  $x$ 's and  $y$ 's to be well mixed in this ordered arrangement. For instance, some of the largest observations should be  $x$ 's, and some  $y$ 's; some of the smallest observations should be  $x$ 's, and some  $y$ 's. However, if the alternative hypothesis that the  $y$ 's are defensive responses is true, then we should expect to find either most of the  $y$ 's being small observations, in which case they would be at the left end of the scale, or most of the  $y$ 's being large observations in which case they would be at the right end of the scale, or a sizable number of  $y$ 's at one end and a sizable number at the other end (some of the defensive responses being "vigilant," others being "repressive"). In any of these three situations the control

group scores will be unduly congested, at one end, at the other end, or in the middle. The character of the test, then, is to determine whether the  $x$ 's (control group scores) are so closely compacted as to call for rejecting the null hypothesis in favor of the alternative hypothesis.

The construction of a suitable test depends, then, on doing two things: 1. Finding a satisfactory statistic,  $s^*$ , to measure the degree to which the  $x$ 's are compacted. 2. Finding a number  $c$  such that if the null hypothesis is true the probability that  $s^*$  should be less than or equal to  $c$  is small, say .05 or .01.

The choice of the statistic  $s^*$  to measure the degree of compactedness of the control group scores must be made in the light of the second requirement; that is, we must be able to find its distribution under the null hypothesis. Accordingly we might take our measure of compactedness to be the range of values occupied by the group of  $x$ 's in the ordered arrangement of all  $x$ 's and  $y$ 's.

Consider the following hypothetical scores:

$x$ (controls)		$y$ (experimentals)	
1.7	2.7	1.3	3.5
1.8	2.8	1.5	3.8
1.9	3.1	1.6	4.5
2.2	3.6	3.4	6.7
	3.7		

We arrange these in order of increasing size and obtain the array:

1.3	1.5	1.6	1.7	1.8	1.9	2.2	2.7	2.8	3.1	3.4	3.5	3.6	3.7	3.8	4.5	6.7
$y$	$y$	$y$	$x$	$x$	$x$	$x$	$x$	$x$	$x$	$y$	$y$	$x$	$x$	$y$	$y$	$y$
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

In this array the obtained scores are given in the first row, their identifications as  $x$  or  $y$  in the second row, and their ranks (from least to greatest) in the third row. The smallest  $x$  has rank 4, the largest  $x$  has rank 14. The total "span" of the  $x$ 's is over 11 ranks; that is, in order to include all the  $x$ 's we must take a consecutive block of 11 ranks, 4 through 14 inclusive. We shall define our measure of compactedness to be the smallest number of consecutive ranks necessary to include all the  $x$ 's; we shall use the symbol  $s^*$  to denote this statistic. Observe that  $s^*$  is equal to one plus the difference between the extreme  $x$  ranks. For example,  $s^* = 11$ ,  $m = 9$ ,  $n = 8$ . It would remain to determine whether 11 is too small a value for  $s^*$  reasonably to arise by chance if the  $x$ 's and  $y$ 's are from a common population.

A defect of taking the measure of compactedness to be  $s^*$ , as defined, is that it depends so strongly on the two extreme scores of the control group. This is open to objections; first, if the control group be contaminated with a few members who might more properly be classified as experimentals, often

when we should reject the null hypothesis, we will fail to do so merely because these few members of the control group will by themselves determine the value of the statistic  $s^*$ ; second, where  $m$  is large, the range (obviously about the same thing as "span") is an inefficient index of the compactedness or spread of the group of  $x$ 's.

Both objections can be met in part by considering modifications of  $s^*$ . We might instead of  $s^*$  consider the span of the middle  $m - 2h$   $x$ 's where  $h$  is arbitrarily selected as some fairly small number. That is, we find the span of the  $m$   $x$ 's after the extreme  $h$   $x$ 's at each end are omitted from consideration. For instance, in the example if we should decide to take  $h = 2$ , then  $m - 2h = 9 - 4 = 5$ . These 5 middle  $x$ 's occupy ranks 6 through 10 and this "truncated span" is 5. We can denote this by the statement

$$s_h^* = 5, \quad h = 2.$$

So, we define  $s_h^*$  as the least number of consecutive ranks necessary to include all the  $x$ 's, except for the  $h$  least and the  $h$  greatest of them. That is, the  $h$  least and  $h$  largest  $x$ 's are omitted from consideration, and  $s_h^*$  is the span of the remaining  $m - 2h$   $x$ 's.

If we adopt  $s_h^*$  as our test statistic we then have the task of finding the probabilities associated with small values of  $s_h^*$ . Since  $s_h^*$  can never be less than  $m - 2h$ , and can never be greater than  $n + m - 2h$ , the problem is to determine how much larger than  $m - 2h$  it can become and still remain convincingly small. What is needed, then, is to find the largest possible integer  $r$ , such that for a predetermined level of significance,  $\alpha$ ,

$$P\{s_h^* \leq m - 2h + r\} \leq \alpha. \quad (1)$$

When the largest such value of  $r$  is found then we have the criterion: Reject the null hypothesis if

$$s_h^* \leq m - 2h + r \quad (2)$$

and otherwise accept it. The distribution theory of  $s_h^*$  is worked out in the brief mathematical appendix. The result is given in formula (3), which states that where there are  $m$   $x$ 's and  $n$   $y$ 's and  $s_h^*$  is defined as above, the probability that  $s_h^* \leq m - 2h + r$  is equal to the sum of  $r + 1$  terms, each a product of certain combinatorials, divided by the total possible combinations of  $m + n$  objects taken  $m$  at a time.

$$P\{s_h^* \leq m - 2h + r\} = \frac{\sum_{i=0}^r \binom{i + m - 2h - 2}{i} \binom{n + 2h + 1 - i}{n - i}}{\binom{m + n}{m}}, \quad (3)$$

where for any positive integers,  $a$  and  $b$ , the symbol  $\binom{a}{b}$  represents the combinations of  $a$  things taken  $b$  at a time; that is,

$$\binom{a}{b} = \frac{a!}{(a-b)!b!} \quad \text{if } a \geq b, \\ = 0 \quad \text{if } a < b. \quad (4)$$

Then for any  $m$ ,  $n$ , and  $h$  one can find the largest value of  $s_h^*$  which is significant at some significance level, say .05, by cumulating the terms in the numerator until reaching the largest value of  $r$  for which the quotient is smaller than .05. A table of significant values of  $s_h^*$  for various  $m$ ,  $n$ ,  $h$  is given; for other values of  $m$  and  $n$  or  $h$  the investigator may find significant values by using formula (3).

For any choice of  $h$  between zero and  $(m-1)/2$  there is a test based on the corresponding  $s_h^*$ . Only one such test should be applied to any one set of data. Which  $h$  to choose is a matter not too easily decided. Any choice will be legitimate so long as it is made before the observations are taken. In general as  $m$  increases,  $h$  should increase. The tables given in this paper have taken  $h = m/4$ . A relatively large value of  $h$  such as this protects the experimenter if the control group is contaminated with subjects who should better have been in the experimental group. A small value of  $h$ , such as say  $h = m/15$  or  $h = 0$ , protects the experimenter if the experimental group is contaminated with subjects who should better have been in the control group.

*Example:*

Suppose that there are 16 in the control group ( $m = 16$ ) and 12 in the experimental group ( $n = 12$ ) and that the scores are:

$x$ :	21.0,	21.4,	22.5,	23.6,	25.8,	25.9,	27.0,	27.8
	30.1,	30.5,	33.7,	34.8,	35.1,	35.3,	38.9,	45.0
$y$ :	17.4,	18.6,	19.1,	22.3,	24.7,	25.2		
	26.1,	27.1,	28.0,	29.5,	30.3,	30.4		

Arrangement of these scores from least to greatest yields the following sequence of  $x$ 's and  $y$ 's, together with their ranks.

$y \ y \ y \ x \ x \ y \ x \ x \ y \ y \ x \ x \ y \ x \ y \ x \ y \ y \ x \ y \ y \ x \ x \ x \ x \ x \ x \ x \ x$   
 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

In this array the values of  $s_h^*$  for certain values of  $h$  are

$$\begin{aligned} h = 0; s_h^* &= 28 - 4 + 1 = 25 \\ h = 1; s_h^* &= 27 - 5 + 1 = 23 \\ h = 2; s_h^* &= 26 - 7 + 1 = 20 \\ h = 3; s_h^* &= 25 - 8 + 1 = 18 \\ h = 4; s_h^* &= 24 - 11 + 1 = 14 \end{aligned} \quad (5)$$

For  $m = 16$ ,  $n = 12$ , and  $h = 4$  we find from Table 2 that  $s_h^*$  as small as 10 is significant at level .057. Since our value 14 is not this small we find

no evidence of defensive behavior in the experimental group, as compared with the control group.

Had the experimenter, *in advance of taking the observations*, decided to use  $h = 1$ , then the value of his test statistic would be 23 and he would have to find the significance level for the statistic by using formula (3). In doing

TABLE 1  
Probability That  $s_h^* \leq m + r - 2h$   
When  $m = 12$ ,  $h = 3$ , for Various Values of  $n$

$n$	$m + r - 2h$			
	6	7	8	9
6	.092			
8	.051			
12	.019	.077		
18	.006	.026	.068	
24	.002	.010	.029	.062

TABLE 2  
Probability That  $s_h^* \leq m + r - 2h$   
When  $m = 16$ ,  $h = 4$ , for Various Values of  $n$

$n$	$m + r - 2h$						
	8	9	10	11	12	13	14
8	.033	.142					
12	.010	.048	.133				
16	.003	.019	.057				
24	.001	.004	.013	.032	.065		
32	.000	.001	.004	.010	.021	.041	.070

TABLE 3  
Probability That  $s_h^* \leq m + r - 2h$   
When  $m = 20$ ,  $h = 5$ , for Various Values of  $n$

$n$	$m + r - 2h$									
	10	11	12	13	14	15	16	17	18	19
10	.012	.062								
15	.002	.015	.049	.118						
20	.001	.004	.015	.041	.088					
30	.000	.001	.002	.006	.015	.031	.058			
40	.000	.000	.000	.001	.003	.007	.015	.026	.044	.069

TABLE 4  
Probability That  $s_h^* \leq m + r - 2h$   
When  $m = 32$ ,  $h = 8$ , for Various Values of  $n$

n	$m + r - 2h$															
	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
16	.001	.004	.018	.055												
24	.000	.000	.002	.006	.017	.040	.081									
32	.000	.000	.000	.001	.002	.007	.016	.032	.059							
48	.000	.000	.000	.000	.000	.000	.001	.002	.004	.008	.014	.023	.037	.058		
64	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001	.001	.003	.005	.008	.012	.019

this computation the work would be greatly facilitated by use of a table of logarithms of factorials.

In case there are ties in the ranks the value of  $s_h^*$  may be different for different ways of breaking the ties. When this is true break the ties in *all possible ways*, find the probability level associated with the sample for each such way, and average these probabilities.

#### Appendix

Where there are  $m$   $x$ 's and  $n$   $y$ 's arranged in order of ascending size define  $s_h^*$  as the least number of consecutive ranks necessary to include all  $x$ 's except the first  $h$  of them and the last  $h$  of them.

Then

$$m - 2h \leq s_h^* \leq m + n - 2h. \quad (6)$$

The problem is to find  $P\{s_h^* = q\}$ ,  $m - 2h \leq q \leq m + n - 2h$ .

To find the probability that  $s_h^* = q$  we first find the number of ways in which we can have  $s_h^* = q$ , and then divide by the total number of ways in which we can fill the  $m + n$  rank positions with  $m$   $x$ 's and  $n$   $y$ 's.

To have  $s_h^* = q$  we must have  $h$   $x$ 's in the first  $j$  observations, then an  $x$ , then  $m - 2h - 2$   $x$ 's in the next  $q - 2$  observations, then an  $x$ , then  $h$   $x$ 's in the last  $m + n - j - q$  observations. Here  $h \leq j \leq m + n - q - h$ . Diagrammatically this statement is represented in Figure 1.

Here the least  $x$  (except for the  $h$  smallest of them) occurs at the left-hand encircled rank. To the left there are  $j$  smaller observations, of which exactly  $h$  are  $x$ 's. The right-hand encircled rank represents the largest  $x$

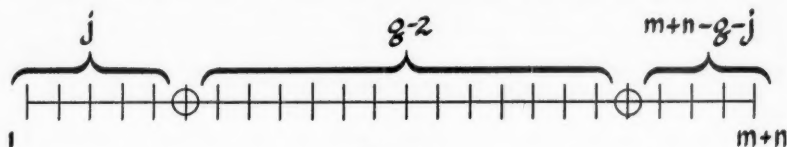


FIGURE 1

(except for the largest  $h$  of them). Between the two encircled ranks lie  $q - 2$  observations, among which are all the  $x$ 's except for the  $h$  largest,  $h$  smallest, and the two  $x$ 's at the ends of span  $s_h^*$ ; that is, there are  $m - 2h - 2$   $x$ 's among these  $q - 2$  observations. The remaining  $m + n - q - j$  observations to the right of the right-hand encircled rank contain exactly  $h$   $x$ 's. Then for any fixed  $j$  between  $h$  and  $n + h$ , inclusive, the number of ways in which we can have  $s_h^* = q$  is

$$\binom{j}{h} \binom{q-2}{m-2h-2} \binom{m+n-q-j}{h}. \quad (7)$$

And to find the total number of ways in which we can have  $s_h^* = q$  we must sum this over the range of  $j$ , giving as the number of ways for  $s_h^* = q$ , the quantity

$$\sum_{j=h}^{m+n-q-h} \binom{j}{h} \binom{q-2}{m-2h-2} \binom{m+n-q-j}{h}, \quad (8)$$

which may be written (by replacing  $j$  by  $k + h = j - h$ ) as:

$$\binom{q-2}{m-2h-2} \sum_{k=0}^{m+n-q-2h} \binom{k+h}{h} \binom{m+n-q-k-h}{h}. \quad (9)$$

To evaluate this we use an identity of probability theory,\*

$$\sum_{j=0}^K \binom{B+J-1}{J} \binom{A+K-J-1}{K-J} = \binom{A+B+K-1}{K}. \quad (10)$$

Taking  $h + 1$  for  $B$ ,  $m + n - 2h - q$  for  $K$ ,  $h$  for  $A - 1$ , and  $k$  for  $J$ , we obtain

$$\binom{q-2}{m-2h-2} \binom{m+n-q+1}{m+n-q-2h}. \quad (11)$$

Now in order to find the number of ways in which we can have  $s_h^* \leq m - 2h + r$ , we must add the above expression for values of  $q$  less than or equal to  $m - 2h + r$ . Recalling that necessarily  $q \geq m - 2h$ , this gives the expression

$$\sum_{q=m-2h}^{m-2h+r} \binom{q-2}{m-2-2h} \binom{m+n-q+1}{m+n-q-2h}, \quad (12)$$

which may be written (by replacing  $q$  by  $i + m - 2h$ ) as

$$\begin{aligned} \sum_{i=0}^r \binom{i+m-2h-2}{m-2h-2} \binom{n+2h+1-i}{n-i} \\ = \sum_{i=0}^r \binom{i+m-2h-2}{i} \binom{n+2h+1-i}{n-i}. \end{aligned} \quad (13)$$

\*Feller, W. *An Introduction to Probability Theory and Its Applications*. New York: John Wiley & Sons, 1950, p. 48.

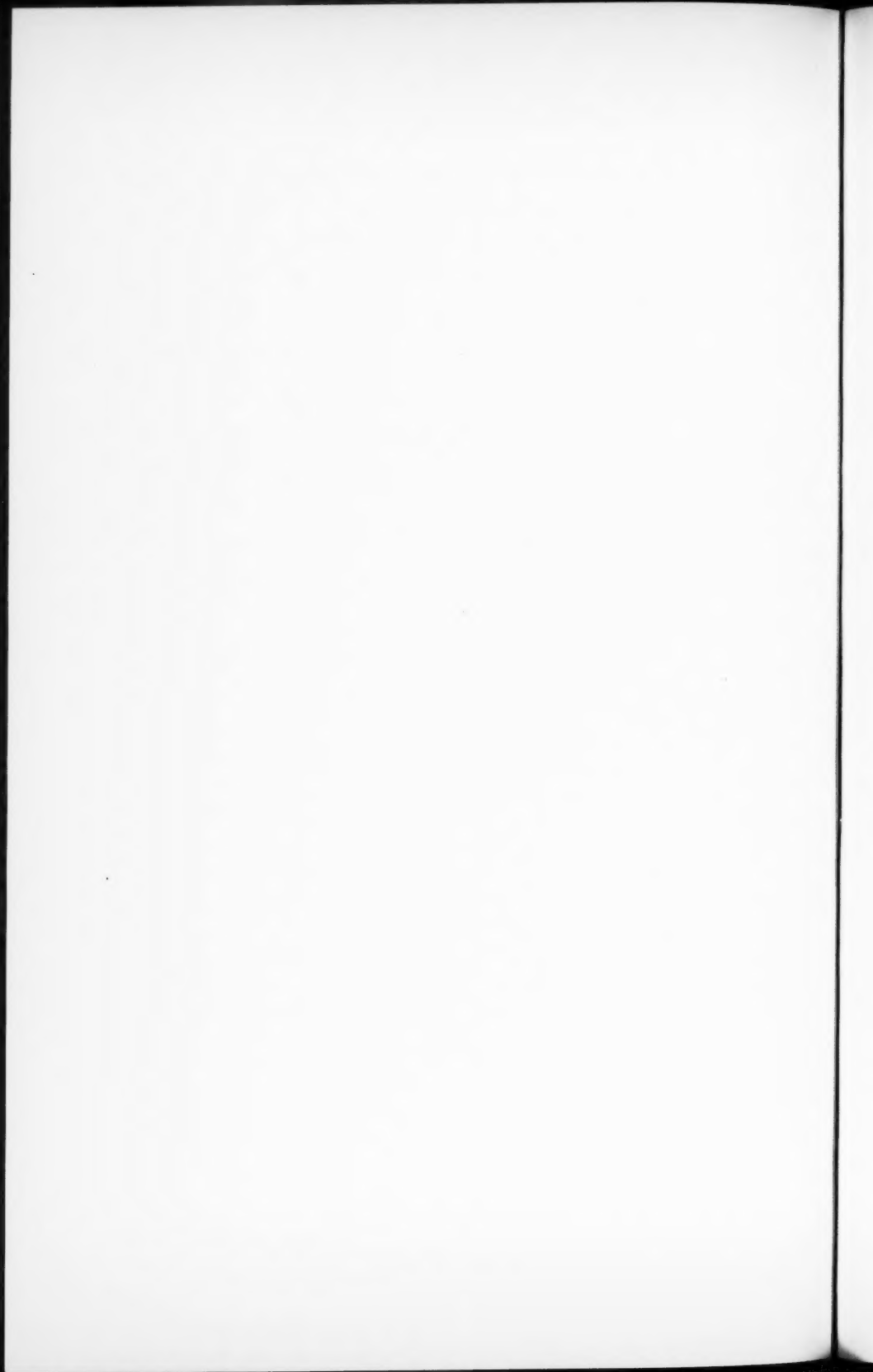
Application of the identity earlier cited shows that when  $r = n$ , that is, we sum over all possible values of  $s_h^*$ , the sum becomes  $\binom{m+n}{m}$ , as it should. Therefore, to write the probability that  $i \leq r$  or that  $q \leq m - 2h + r$ , which is the same thing, we divide the above sum by  $\binom{m+n}{m}$  and obtain

$$P\{s_h^* \leq m - 2h + r\} = \frac{\sum_{i=0}^r \binom{i+m-2h-2}{i} \binom{n+2h+1-i}{n-i}}{\binom{m+n}{m}}. \quad (14)$$

And thus the correctness of formula (3) is demonstrated. It may be remarked that the distribution of  $s_h^*$  is asymptotically normal, but the approximation is unsatisfactory for sample sizes of practical interest.

*Manuscript received 12/6/51*

*Revised manuscript received 2/29/52*



## A NOTE ON CLUSTER-DIRECTED ANALYSIS

C. J. ADCOCK

VICTORIA UNIVERSITY COLLEGE

Thurstone's multiple group method of factor analysis has been widely used as a basis for rotation to simple structure. To make the most of the economy offered by this method, factor axes may be located directly by the correlation clusters; and methods of doing this are here discussed.

Thurstone's multiple group method (1) of factoring provides a simple and rapid means of analysis, but it depends upon rotation to simple structure to fix the reference axes. Since cluster selection can itself be used as a means to locate the reference axes (2, 3), it may be desirable to take advantage of the group method to avoid the tedious and sometimes uncertain task of rotation according to simple structural principles. This will involve more careful selection of clusters than is demanded in the usual application of the Thurstone group method; but, in most cases, the work will be considerably less than would be involved in the normal process of rotation.

The objections to such a procedure as this are two. In the first place, factors based on clusters of the original correlations show considerable inter-correlation in most cases. For this reason it is usual to first remove the general factor either as a first centroid (2) or as a "basic" factor (3). In the second place, rotation to orthogonality is needed to check on the residual correlation. It should be noted, however, that oblique factors plus a second-order factor which accounts for this inter-correlation can be converted into equivalent orthogonal factors plus a general factor (4, p. 297) and this is what is achieved by Burt's group method. Burt's method, however, makes no provision for overlap, and he has to deal with such a situation by the somewhat drastic method of carrying out an ordinary centroid analysis in addition and then rotating to the group-factor positions. Actually, he carries out the centroid analysis first and bases his matrix sections on this instead of the cluster analysis. In practice this double analysis is much shorter than the usual rotation methods, as was demonstrated by Charlotte Banks (5); but the cluster-directed approach is still shorter and may have special advantages with some data.

If we are prepared to accept orthogonal factors plus a general or basic factor instead of the usual oblique factors plus second-order factors (and on the grounds of simplicity a strong argument can be advanced in favor of this), we may meet both the above objections to cluster-based rotation of the

Thurstone's group-factor matrix at once by an alternative method of rotating to orthogonality. Any orthogonal rotation will, of course, remove the inter-correlation but Thurstone's method of basing on a diagonal analysis will not serve our purpose, since the first factor is used as a pivot and is so arbitrarily weighted. Subsequent factors are increasingly distorted and, if many factors are involved, the last factor may be forced far from the position suggested by its correlation cluster.

We require an alternative method of rotation which will keep as closely as possible to the positions suggested by all the clusters. Under these circumstances, provided we have made an adequate cluster search, we should find that one of our orthogonal factors corresponds to the basic factor of Burt, since it too must have a cluster effect. We can thus get a complete system of orthogonal factors by one simple rotation of the group-factor matrix.

The alternative method of rotation to orthogonal position here referred to is one which requires that all the reference axes participate equally in the "give and take" required to attain orthogonality. This means that the transformation matrix,  $F_{pm}^{-1}$ , must be symmetric. Our problem is thus to calculate from the interfactor correlation matrix,  $R_{pq}$ , a factor matrix  $F_{pm}$  such that  $F_{pm} = F'_{pm}$ . I am indebted to Dr. Ledyard Tucker for a precise solution to this problem. He takes as his basic equation,

$$R_{pq} = \Lambda D \Lambda' \text{ (where } \Lambda \text{ is orthogonal).} \quad (1)$$

This is equation 110 from Thurstone's text (1, p. 501). It applies to the principal-axes solution, the entries in the diagonal matrix  $D$  being the characteristic roots of  $R_{pq}$ . It is required that

$$F_{pm} = F'_{pm}. \quad (2)$$

If we take

$$F_{pm} = \Lambda D^{\frac{1}{2}} \Lambda', \quad (3)$$

which is clearly symmetric, then

$$\begin{aligned} F_{pm} F'_{pm} &= \Lambda D^{\frac{1}{2}} \Lambda' \Lambda D^{\frac{1}{2}} \Lambda' \\ &= \Lambda D^{\frac{1}{2}} I D^{\frac{1}{2}} \Lambda' \\ &= \Lambda D \Lambda' \\ &= R_{pq} \quad \text{by (1) as required} \end{aligned} \quad (4)$$

and

$$F_{pm}^{-1} = \Lambda D^{-\frac{1}{2}} \Lambda'. \quad (5)$$

If we denote the factor matrix for the principal axes  $F_{ax}$ , we get, in terms of Thurstone's equation 105 (1, p. 501),

$$\Lambda = F_{ax} D^{-\frac{1}{2}}. \quad (6)$$

Substituting in (5) we get

$$\begin{aligned} F_{pm}^{-1} &= (F_{ax} D^{-1}) D^{-1} (F_{ax} D^{-1})' \\ &= F_{ax} D^{-1} D^{-1} F_{ax}' \end{aligned}$$

Post-multiplication of the original oblique factor matrix by  $F_{pm}^{-1}$  gives us the required orthogonal transformation,

$$V F_{pm}^{-1} = F.$$

Once the principal-axes solution has been obtained, the transformation matrix is quickly obtained by equation (7). The chief objection to be urged against the method is the length of time required to find the principal axes, but where only a few factors are concerned this may not be too time-consuming. For correlation tables involving high error variance, I have made use of an approximation method which involves considerably less work. Since much material hardly warrants laborious calculations and since the high error variance often results in considerable distortion when the usual simple-structure methods of rotation are employed, it may be worth while to describe this simpler method.

Briefly, the procedure is to obtain a transformation matrix as described by Thurstone and then to calculate a correction matrix which will rotate this to symmetric form (i.e.,  $F_{pm}^{-1}$  as above). The correction matrix should be orthogonal in order to keep the factor variance constant; and if the non-diagonal terms are chosen skew-symmetrically, the solution of the equations involved is facilitated. These two requirements are to some extent conflicting, but if the matrix is first calculated according to the latter requirement and then normalized, the compromise seems sufficiently close for the purpose here described. It should be noted that the requirement of skew symmetry is not exactly true of the precise solution. The skew symmetric terms can be calculated one at a time if we take the diagonal elements as unity. Denote the elements of the original transformation matrix  $t_{ij}$ , the elements of the correction matrix  $c_{ij}$ , and the elements of  $F_{pm}^{-1}$  as  $f_{ij}$ . Now since  $F_{pm}^{-1}$  is symmetric,  $f_{23} = f_{32}$ ; so for a  $3 \times 3$  matrix we can write

$$t_{22}c_{23} + t_{23}c_{33} = t_{33}c_{32}.$$

But  $c_{33} = \text{unity}$  and  $c_{32} = -c_{23}$ . Substituting, we get

$$t_{22}c_{23} + t_{23} = -t_{33}c_{23}, \quad (8)$$

from which

$$c_{23} = \frac{-t_{23}}{t_{22} + t_{33}};$$

and in the same way we find

$$c_{13} = \frac{-(t_{13} + t_{12}c_{23})}{t_{11} + t_{33}}, \quad (10)$$

$$c_{12} = \frac{t_{13}c_{23} - t_{12} - t_{23}c_{13}}{t_{11} + t_{22}}. \quad (11)$$

Proceeding in this way we can deal with matrices of any size, calculating the entries above the diagonal and simply changing the signs for the corresponding entries below the diagonal. We append the results obtained by the precise and approximate methods for a simple example. It will be seen that there is an error of the order of .02, which is quite as good an agreement as we are likely to get from two simple-structure rotations.

TABLE 1  
Approximate Solution

1.00	-.12	-.11	
	1.01	-.71	$T_{ij}$
		1.23	
By formulas:			
1.000	.066	.066	
-.066	1.000	.317	$C_{ij}$
-.066	-.317	1.000	
Normalizing:			
.991	.063	.063	
-.065	.951	.302	$C_{ij}D$
-.065	.302	.951	
Symmetric transformation matrix:			
1.006	-.018	-.078	
-.019	1.175	-.370	$T_{ij}C_{ij}D = F^{-1}$
-.080	-.371	1.170	

TABLE 2  
Precise Solution

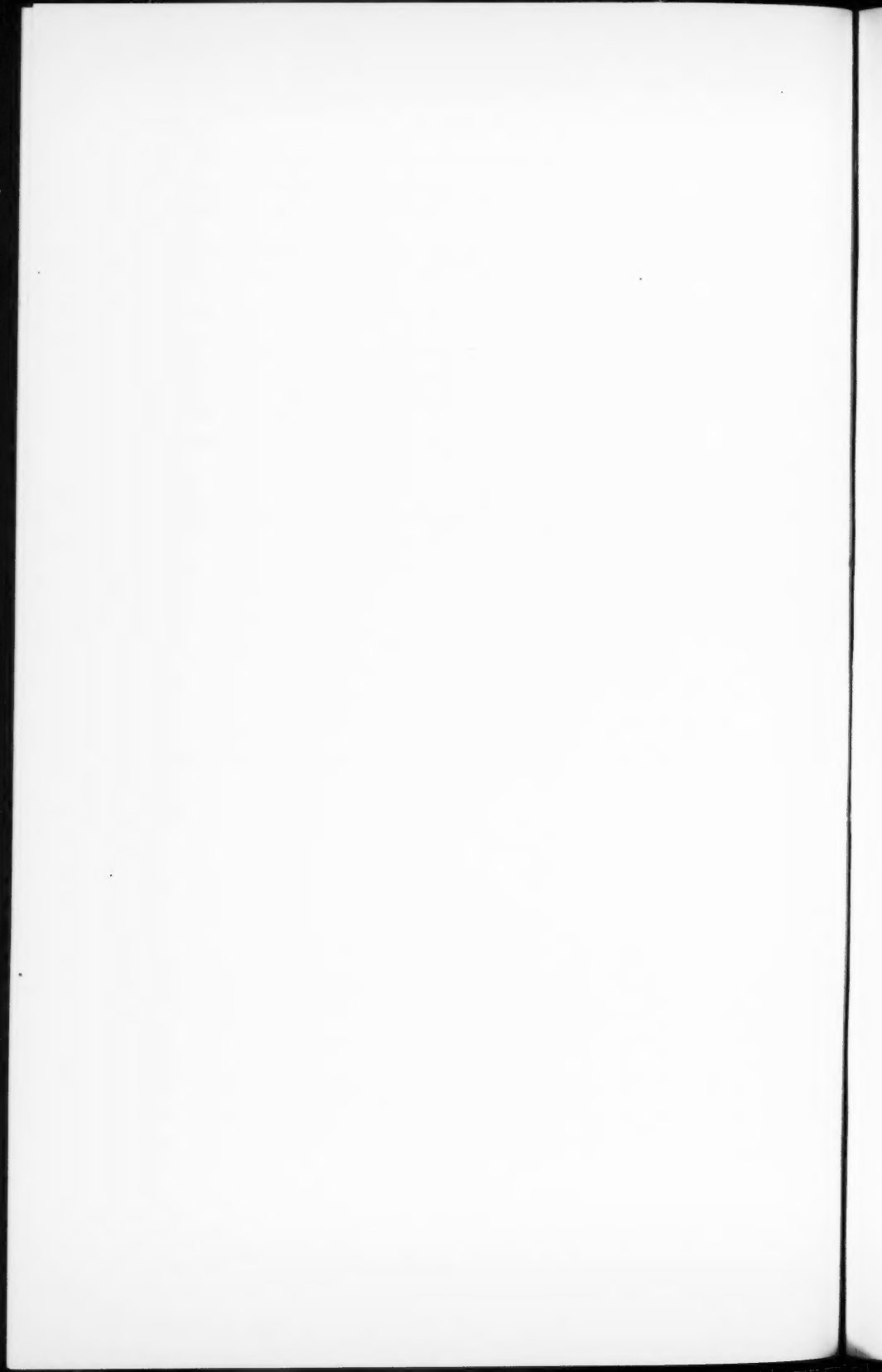
Original interfactor Correlation:			
1.00	.12	.16	
.12	1.00	.58	$R_{pq}$
.16	.58	1.00	
Symmetric transformation matrix:			
1.008	-.039	-.062	
-.039	1.174	-.373	$F^{-1}_{pm}$
-.062	-.373	1.176	

## REFERENCES

1. Thurstone, L. L. Multiple-factor analysis. Chicago: Univ. Chicago Press, 1947.
2. Tucker, L. R. A semi-analytic method of factorial rotation to simple structure. *Psychometrika*, 1944, 9, 43-68.
3. Burt, C. Group factor analysis. *Brit. J. Psychol. (Statist. Sect.)*, 1950, 3, 40-75.
4. Thomson, G. H. The factorial analysis of human ability. London: Univ. London Press, 1939.
5. Banks, C. Primary personality factors in women: a re-analysis. *Brit. J. Psychol. (Statist. Sect.)*, 1948, 1, 204-218.

*Manuscript received 7/4/50*

*Revised manuscript received 12/7/51*



# METHODS OF SOLVING SOME PERSONNEL-CLASSIFICATION PROBLEMS\*

D. F. VOTAW, JR.

YALE UNIVERSITY

The personnel-classification problems considered in this paper are related to those studied by Brogden (2), Lord (6), and Thorndike (8). Section 1 gives an approach to personnel classification. A basic problem and variations of it are treated in section 2; and the computation of a solution is illustrated in section 3. Two extensions of the basic problem are presented in section 4. Most of the methods indicated for computing solutions are applications of the "simplex" method used in linear programming (see 1, Chs. XXII, XXIII). The capabilities of a high speed computer in regard to the simplex method are discussed briefly (see section 1).

1. *Introduction.* The type of problem to be considered is indicated by the illustration below. Consider two persons and two jobs, and suppose that the productivity of each person with regard to each job is known. Suppose further that each job must be filled and that each person must be assigned to one and only one job. It is assumed that the productivities can be represented as single numbers. For example, let them be as follows:

		Jobs	
		1	2
Persons	1	9	7
	2	8	5

(1:1)

Thus 9 is the productivity of person 1 on job 1, 7 is the productivity of person 1 on job 2, etc. To make maximum use of manpower the assigning agency would wish to establish an order of preference for the possible assignments and choose the assignment that heads the order. It would be natural to base such an ordering on average productivity. In the case at hand there are

\*This work was supported in part by the United States Air Force under Contract AF 33(038)-25192, monitored by the Personnel Research Laboratory of the Air Training Command Human Resources Research Center. Permission is granted for reproduction, translation, publication, use, and disposal in whole or in part by or for the United States Government.

two possible assignments. From (1:1) we obtain the following averages of the productivities associated with them:

Assignment	Average Productivity	
Person 1 on Job 1, Person 2 on Job 2	$(9 + 5)/2 = 7$	
Person 1 on Job 2, Person 2 on Job 1	$(7 + 8)/2 = 7.5$	(1:2)

It follows that the preferred assignment is formed by placing person 1 on job 2 and person 2 on job 1. This gives the maximum possible average, which is 7.5. It should be noted that comparing assignments on the basis of the average productivity is equivalent to comparing them on the basis of the sum of productivities (see (1:2)).

The situation described above involving two persons and two jobs suggests the following general problem: Given  $N$  persons,  $N$  jobs, and the productivity of each person on each job; find an assignment of persons to jobs such that the average productivity is a maximum. (Recall that it is assumed that each job must be filled and that each person must be assigned to one and only one job.)\* Problems similar to this have been considered by Brogden (2), Lord (6), and Thorndike (8, p. 217). A basic form of the problem and certain modifications of it are discussed in section 2. Section 3 gives an illustration of a step-by-step computation of a solution. Extensions of the basic problem are formulated in section 4.

Occasionally in personnel classification (e.g., in the military) there may be fewer job categories than jobs and/or fewer personnel categories than persons. This possibility has been taken into account in section 2 in the formulation of the basic problem. The  $N$  persons are considered to be such that  $a_1$  are alike,  $a_2$  are alike,  $\dots$ ,  $a_m$  are alike ( $\sum_{i=1}^m a_i = N$ ); thus there are only  $m$  distinct personnel categories. Similarly, regarding the  $N$  jobs it is assumed that  $b_1$  are alike,  $b_2$  are alike,  $\dots$ ,  $b_n$  are alike ( $\sum_{i=1}^n b_i = N$ ); thus there are only  $n$  distinct job categories. ( $m = n = N = 2$  and  $a_1 = a_2 = b_1 = b_2 = 1$  in the illustration associated with (1:1)). If, say,  $x_{ij}$  persons of type  $i$  are placed on jobs of type  $j$ , their contribution to the "group productivity" is assumed to be  $c_{ij}x_{ij}$ , where  $c_{ij}$  denotes the productivity (or expected productivity) of the  $i$ th type of person on the  $j$ th type of job ( $i = 1, \dots, m; j = 1, \dots, n$ ). For a discussion of how the  $c_{ij}$ 's might be determined see (2, 8); in some situations  $c_{ij}$  might be considered as a monetary saving.

Two extensions of the basic problem will be considered (see section 4). In one of them the number,  $m$ , of personnel categories is assumed to be indefinitely large, but the number,  $n$ , of job categories is assumed to be finite. Each personnel category is represented as a point in  $n$ -space, where the  $n$  coordinates of the point represent the  $n$  productivities associated with the category. An  $n$ -dimensional distribution is assumed to be associated with the

\*In an actual situation the numbers of persons and jobs might be unequal; however, equality could be forced by introducing "dummy" jobs or persons.

set of personnel categories. The essential characteristic of the second extension is that a vector  $c_{1,ij}, \dots, c_{k,ij}$  replaces the quantity  $c_{i,j}$ . This extension might be of interest when, say, skill and preference of each person regarding each job must be considered separately; e.g.,  $c_{1,ij}$  could represent skill and  $c_{2,ij}$  could represent preference (here  $k = 2$ ).

Most of the computational techniques suggested in this paper originated in linear programming. Hand computation of the solution of a general linear programming problem is illustrated in 1, Chapter XXII. It seems possible that an "analogy" machine for solving such problems could be developed (see 3, p. 73). The example in section 3 illustrates hand computation of a solution of the special linear programming problem referred to in section 2 (see also 1, Chapter XXIII). The National Bureau of Standards Eastern Automatic Computer (known as the SEAC) is coded for carrying out the simplex method used in the illustration. The SEAC is capable of solving a personnel-classification problem within a few hours when  $m + n \leq 60$ . It has obtained solutions of problems in which  $m = 4$ ,  $n = 9$ , and  $m = n = 10$  in six minutes and twenty minutes, respectively. Government and military agencies have access to the SEAC through the National Bureau of Standards.

2. *A Basic Problem in Personnel Classification.* Suppose that with regard to  $N$  persons there are  $m$  mutually exclusive personnel categories, and that with regard to  $N$  jobs there are  $n$  mutually exclusive job categories. Let  $a_i$  and  $b_j$  be, respectively, the number of persons in the  $i$ th category and the number of jobs in the  $j$ th category ( $i = 1, \dots, m; j = 1, \dots, n$ ). Assume that any two persons in the same category are essentially identical and that any two jobs in the same category are essentially identical. (This assumption involves no loss of generality since the case in which a job category or personnel category contains only one member is not excluded.) Let  $c_{ij}$  be the productivity (or expected productivity) of any person in the  $i$ th category on any job in the  $j$ th category. Let  $x_{ij}$  represent any number of persons in the  $i$ th category who can be placed on jobs in the  $j$ th category; the array  $(x_{ij})$  is termed an "allocation." Let  $T = \sum_{i,j} c_{ij}x_{ij}$ , which might be considered as the group productivity (or expected group productivity) associated with  $(x_{ij})$ . It is required to find an allocation for which  $T$  is maximized (see 8, p. 217). More specifically, the problem\* is to find values  $x_{11}^{(0)}, \dots, x_{mn}^{(0)}$ , of  $x_{11}, \dots, x_{mn}$ , respectively, for which  $T$  assumes its maximum value, subject to the conditions:

$$\begin{aligned} \sum_i x_{ij} &= b_j, \\ (i &= 1, \dots, m; j = 1, \dots, n), \\ \sum_j x_{ij} &= a_i, \end{aligned} \quad (2:1)$$

\*In regard to this problem the author wishes to acknowledge valuable discussions with S. Kakutani, J. W. Tukey, M. A. Woodbury, and J. T. Dailey.

[in (2:1) each  $a_i$ ,  $b_i$ , and  $x_{ij}$  is a non-negative integer and  $\sum_i a_i = \sum_i b_i = N$ ]. This can be shown to be equivalent (except in minor respects) to the Hitchcock-Koopmans transportation problem, a special topic in linear programming, which can be solved by means of the simplex method (see 1, Chapter XXIII). Professor J. von Neumann has shown\* that this problem is essentially the same as that of finding a best strategy in a certain zero-sum two-person game having a payoff matrix of order  $(m + n) \times mn$ . (See 7.)

An interesting special case of the problem is that in which each  $c_{ij}$  has only two possible values,† say 1 and 0 (which, for example, could mean "qualified" and "not qualified," respectively). A variant of this special case is to determine whether  $N$  is the maximum of  $T$ —and, if so, to find an allocation for which  $T = N$ . (If the maximum of  $T$  is less than  $N$ , presumably the assigning agency would change one or more of the numbers  $a_1, \dots, a_m, b_1, \dots, b_n$  or lower the standards regarding qualifications of persons for jobs.) The special case and this variant can be solved by means of Theorem 3 in (5).

A problem entirely similar to the one stated above (2:1) arises when  $c_{ij}$  is considered not as a productivity but as the cost of training any person in the  $i$ th category to do any job in the  $j$ th category. In this situation  $T$  represents the total cost of training, and one wishes to find an allocation for which  $T$  assumes its *minimum* value.

When  $N$  is large, one might wish to consider the numbers  $a_1, \dots, a_m, b_1, \dots, b_n, x_{11}, \dots, x_{mn}$  as population proportions; the conditions stated in (2:1) would then be replaced by the following conditions:

$$\begin{aligned}\sum_i x_{ij} &= b_j, \\ \sum_j x_{ij} &= a_i, \\ x_{ij} &\geq 0,\end{aligned}\tag{2:2}$$

( $\sum_i a_i = \sum_i b_i = 1$ ;  $a_i, b_i > 0$ ;  $i = 1, \dots, m$ ;  $j = 1, \dots, n$ ). The use of (2:2) in place of (2:1) introduces only minor modifications in the problems stated above.

**3. An Illustrative Example.** In this section a step-by-step procedure for solving problems of the type presented in section 2 will be illustrated. The procedure is an application of the simplex method (1, Chapter XXIII). Each step of the method is characterized by a distinct allocation. Before setting up the illustrative example we shall consider a special property of these allocations and indicate the underlying idea of the method.

A problem is said to be degenerate when a partial sum of  $a$ 's equals a

\*In a lecture at the Princeton University Game Seminar, October, 1951.

†In this case  $m \leq 2^n$ .

partial sum of  $b$ 's (the equality  $\sum_1^m a_i = \sum_1^n b_i$  does not satisfy this condition). (A degenerate problem can be replaced by an essentially equivalent one that is non-degenerate.) In a non-degenerate problem each allocation used in the procedure has exactly  $m + n - 1$  positive  $x_{ij}$ 's. We shall denote them by  $x_{i_1 j_1}, \dots, x_{i_r j_r}$  ( $r = m + n - 1$ ). With this set of  $r$  positive  $x_{ij}$ 's we associate quantities  $u_1, \dots, u_m, v_1, \dots, v_n$ , defined as follows:

$$u_{i_g} + v_{j_g} = c_{i_g j_g} \quad (g = 1, \dots, r). \quad (3:1)$$

These equations can be solved for the  $u$ 's and  $v$ 's, where  $u_1$ , say, is set equal to 0. Let  $u_i + v_j = \bar{c}_{ij}$ , say. The allocation is optimal (i.e.,  $\sum_{i,j} c_{ij} x_{ij}$  assumes its maximum value) if and only if

$$\bar{c}_{ij} \geq c_{ij} \quad (i = 1, \dots, m; j = 1, \dots, n). \quad (3:2)$$

When the allocation associated with a given step is not optimal, one of its positive  $x_{ij}$ 's can be "discarded" and a new positive  $x_{ij}$  can be "introduced" to form an allocation for the next step. The new  $x_{ij}$  can be chosen as one for which the corresponding  $c_{ij}$  yields a largest value of  $c_{ij} - \bar{c}_{ij}$ . This choice determines which one of the positive  $x_{ij}$ 's must be discarded. (See the discussion that follows the computing form below.) Every increase in  $\sum_{i,j} c_{ij} x_{ij}$  from one step to the next is not less than a certain positive number, which depends on the  $c_{ij}$ 's,  $a$ 's, and  $b$ 's. Since the maximum sum is finite, the procedure involves at most a finite number of steps.

The simplex method will be used to solve the problem stated below. Consider a large group of persons and a correspondingly large group of jobs. Assume that there are exactly four types of persons and exactly three types of jobs; thus  $m = 4$  and  $n = 3$ . Suppose that the productivities are as follows:

$$(c_{ij}) = \begin{pmatrix} 9 & 2 & 9 \\ 1 & 8 & 8 \\ 7 & 2 & 1 \\ 9 & 8 & 0 \end{pmatrix} \quad (3:3)$$

( $c_{ij}$  is the actual or expected productivity of a person of type  $i$  on a job of type  $j$ .) Let the proportions of the group associated with the four personnel categories be  $a_1 = .40, a_2 = .20, a_3 = .20, a_4 = .20$ ; and let the proportions of the group required in the three job categories be  $b_1 = .35, b_2 = .35, b_3 = .30$ . (Note that  $\sum_1^4 a_i = \sum_1^3 b_j$ .) The problem is to find an allocation,  $(x_{ij})$ , for which the group productivity,  $\sum_{i,j} c_{ij} x_{ij}$ , assumes its maximum value, where  $(x_{ij})$  is subject to the conditions given in (2:2). The allocation given in

(3:7) is a solution.\* The problem is non-degenerate since no partial sum of  $a$ 's equals a partial sum of  $b$ 's (i.e., it is not the case that an  $a$ , a sum of two  $a$ 's, or a sum of three  $a$ 's equals a  $b$  or a sum of two  $b$ 's).

All allocations used in the procedure have the property that exactly six of the  $x_{ij}$ 's are positive. We may choose any such allocation as one with which to begin the procedure; accordingly, let the first one be as follows:

$$(x_{ij}) = \begin{array}{ccccc} & & & a's & \\ & & & .40 & \\ \left( \begin{array}{ccc} .35 & 0 & .05 \\ 0 & .15 & .05 \\ 0 & 0 & .20 \\ 0 & .20 & 0 \end{array} \right) & & & .20 & \\ & & & .20 & \\ & & & .20 & \end{array} \quad (3:4)$$

$$b's \quad .35 \quad .35 \quad .30 \quad \left( 1.00 = \sum_i a_i = \sum_j b_j \right)$$

The six positive  $x_{ij}$ 's in (3:4) can be considered as having been chosen in the following sequence:  $x_{11}$ ,  $x_{13}$ ,  $x_{42}$ ,  $x_{22}$ ,  $x_{23}$ ,  $x_{33}$ . The way in which each was chosen can be described as follows: choose any  $x_{ij}$ , say  $x_{11}$ , and set  $x_{11}$  equal to its maximum possible value, .35, which is the minimum of  $a_1 = .40$  and  $b_1 = .35$  (this reduces the  $4 \times 3$  array,  $x_{ij}$ , to a  $4 \times 2$  array); next, in the reduced array choose any  $x_{ij}$ , say  $x_{13}$ , and set  $x_{13}$  equal to its maximum possible value, .05, which is the minimum of  $a_1 - x_{11} = .05$  and  $b_3 = .30$ ; continue this method of selection until the original array has been completely reduced.

We now introduce quantities  $u_1, \dots, u_4, v_1, \dots, v_3$  which are to satisfy the following conditions [see (3:1)]:

$$\begin{aligned} u_1 + v_1 &= c_{11}, \\ u_1 + v_3 &= c_{13}, \\ u_4 + v_2 &= c_{42}, \\ u_2 + v_2 &= c_{22}, \\ u_2 + v_3 &= c_{23}, \\ u_3 + v_3 &= c_{33}. \end{aligned} \quad (3:5)$$

The six  $c_{ij}$ 's in (3:5) correspond to the six positive  $x_{ij}$ 's in (3:4). Solving (3:5) for  $u_2, u_3, u_4, v_1, v_2, v_3$  we find that

\*If  $a_1, \dots, a_4, b_1, \dots, b_3$  had been replaced by  $100a_1, \dots, 100a_4, 100b_1, \dots, 100b_3$ , we could substitute the conditions in (2:1) for those in (2:2). A solution of this new problem would then be given by multiplying each  $x_{ij}$  in (3:7) by 100.

$$\begin{aligned}
 v_1 &= c_{11} - u_1, \\
 v_2 &= c_{22} - c_{23} + c_{13} - u_1, \\
 v_3 &= c_{13} - u_1, \\
 u_2 &= c_{23} - c_{13} + u_1, \\
 u_3 &= c_{33} - c_{13} + u_1, \\
 u_4 &= c_{42} - c_{22} + c_{23} - c_{13} + u_1.
 \end{aligned}
 \tag{3.6}$$

For example, from (3.5) we have that  $v_3 = c_{13} - u_1$ ,  $u_3 = c_{33} - v_3 = c_{33} - c_{13} + u_1$ , etc. The allocation in (3.4) is optimal if and only if for every  $(i, j)$   $\bar{c}_{ij} \geq c_{ij}$ , where  $\bar{c}_{ij} = u_i + v_j$  [see (3.2)]. It will be noted from Step 1 on the computing form below that this allocation is not optimal: e.g.,  $\bar{c}_{31} = 1 < c_{31} = 7$ . This requires that a second step of the procedure be performed. The allocation associated with the second step will be formed from (3.4) by discarding  $x_{33}$  and introducing  $x_{31}$  [see the remarks immediately below (3.2)].

The information in (3.6) and the values of the  $\bar{c}_{ij}$ 's are given in Step 1 of the computing form below. The first column of the form gives the positive  $x_{ij}$ 's; the next column gives the values of the corresponding  $c_{ij}$ 's. In the columns headed "Coefficients of  $u_i, \dots, v_j$ " the + 's and - 's given in (3.6) are indicated, where  $u_1$  has been equated to 0. Using the  $c_{ij}$  column and the + 's and - 's in a  $u$  or  $v$  column, we can evaluate the  $u$  or  $v$  rapidly; e.g.,  $u_4 = -9 - 8 + 8 + 8 = -1$  (see the last row of the form for Step 1). The last column gives the products  $c_{ij}x_{ij}$  and their sum. The way of passing from one step to the next is indicated below in discussions of the steps. An interesting feature of the general procedure is that for any step each coefficient in the expression of a  $u$  or  $v$  as a linear combination of  $c$ 's is +1, 0, or -1. The computing form is a slight modification of one given in (4).

*Discussion of Step 1.* It is not the case in Step 1 that each  $\bar{c}_{ij} \geq c_{ij}$ ; thus another step of the procedure must be performed. The maximum of the differences  $c_{ij} - \bar{c}_{ij}$  is  $c_{31} - \bar{c}_{31} = 7 - 1 = 6$ , and so  $x_{31}$  will be given a positive value in the allocation for Step 2 (note that  $\bar{c}_{31}$  has been italicized in Step 1). Since  $\bar{c}_{31} = u_3 + v_1$ , we have that  $\bar{c}_{31} = c_{11} + c_{33} - c_{13}$  (thus  $\bar{c}_{31} + c_{13} = c_{11} + c_{33}$ ). This indicates that by adding, say,  $d > 0$  to  $x_{31}$  and  $x_{13}$  in (3.4) and subtracting  $d$  from  $x_{11}$  and  $x_{33}$  a new allocation will be formed which differs from (3.4) in that the new  $x_{31}$  will be positive and, if  $d$  is sufficiently large, the new  $x_{11}$  or  $x_{33}$  will equal 0. Since in Step 1  $x_{33} = .20$  is smaller than  $x_{11} = .35$ , we set  $d$  equal to .20 (thus the new  $x_{33}$  will be 0). It follows from the expression above for  $\bar{c}_{31}$  that  $c_{33} = \bar{c}_{31} + c_{13} - c_{11}$ ; from this we can easily determine the coefficients of the  $u$ 's and  $v$ 's for Step 2; e.g., in Step 1  $u_3 = c_{33} - c_{13}$ , thus in Step 2  $u_3 = c_{31} + c_{13} - c_{11} - c_{13} = c_{31} - c_{11}$ .

*Discussion of Step 2.* The  $\bar{c}_{ij}$ 's obtained in Step 2 are not such that each

Computing Form for the Simplex Method

Positive $x_{ij}$	$c_{ij}$	Coefficients of						$c_{ij}x_{ij}$
		$u_1$	$u_2$	$u_3$	$u_4$	$v_1$	$v_2$	$v_3$
$x_{11} = .35$	$c_{11} = 9$					+		
$x_{13} = .05$	$c_{13} = 9$		-		-		+	+
$x_{22} = .15$	$c_{22} = 8$				-		+	
$x_{33} = .05$	$c_{33} = 8$		+		+		-	
$x_{33} = .20$	$c_{33} = 1$			+				
$x_{42} = .20$	$c_{42} = 8$				+			
Values of $u$ 's and $v$ 's $\rightarrow$		0	-1	-8	-1	9	9	9
$(c_{ij}) = (u_i + v_j)$ $\begin{pmatrix} 9 & 9 & 9 \\ 8 & 8 & 8 \\ 7 & 1 & 1 \\ 8 & 8 & 8 \end{pmatrix} = \begin{pmatrix} u_1 + v_1 & u_1 + v_2 & u_1 + v_3 \\ u_2 + v_1 & u_2 + v_2 & u_2 + v_3 \\ u_3 + v_1 & u_3 + v_2 & u_3 + v_3 \\ u_4 + v_1 & u_4 + v_2 & u_4 + v_3 \end{pmatrix}$								
								$7.00 = \sum_{i,j} c_{ij}x_{ij}$

Step 1

Step 2

$x_{11} = .15$	$c_{11} = 9$			-		+			1.35
$x_{13} = .25$	$c_{13} = 9$		-		-		+	+	2.25
$x_{22} = .15$	$c_{22} = 8$				-		+		1.20
$x_{23} = .05$	$c_{23} = 8$		+		+		-		.40
$x_{42} = .20$	$c_{42} = 8$				+				1.60
$x_{31} = .20$	$c_{31} = 7$			+					1.40
Values of $u$ 's and $v$ 's $\rightarrow$		0	-1	-2	-1	9	9	9	8.20
									$(c_{ij}) = \begin{pmatrix} 9 & 9 & 9 \\ 8 & 8 & 8 \\ 7 & 7 & 7 \\ 8 & 8 & 8 \end{pmatrix}$
									$= \sum_{i,j} c_{ij}x_{ij}$

Step 2

Step 3

$x_{11} = .10$	$c_{11} = 9$		-	-	-	+	+	+		.90
$x_{13} = .30$	$c_{13} = 9$								+	2.70
$x_{22} = .20$	$c_{22} = 8$		+							1.60
$x_{42} = .15$	$c_{42} = 8$		-					+		1.20
$x_{31} = .20$	$c_{31} = 7$			+						1.40
$x_{41} = .05$	$c_{41} = 9$		+		+			-		.45
Values of $u$ 's and $v$ 's $\rightarrow$		0	0	-2	0	9	8	9		8.25

$$(e_{ij}) = \begin{pmatrix} 9 & 8 & 9 \\ 9 & 8 & 9 \\ 7 & 6 & 7 \end{pmatrix}$$

$$\sum (e_{ij})$$

$$= \sum_{i,j} e_{ij} x_{ij}$$

$\bar{c}_{ij} \geq c_{ij}$ ; hence a third step of the procedure must be performed. There is one (and only one) case in which  $c_{ij} - \bar{c}_{ij}$  is positive—namely,  $c_{41} - \bar{c}_{41} = 9 - 8 = 1$ ; thus  $x_{41}$  should be positive in the new allocation. (Note that  $\bar{c}_{41}$  is italicized in Step 2.) Since  $\bar{c}_{41} + c_{13} + c_{22} = c_{11} + c_{23} + c_{42}$ , it follows that .05 can be added to  $x_{41}$ ,  $x_{13}$ , and  $x_{22}$  and subtracted from  $x_{11}$ ,  $x_{23}$ , and  $x_{42}$  ( $x_{23}$ 's value becomes 0). It should be noted that  $c_{23} = \bar{c}_{41} + c_{13} + c_{22} - c_{11} - c_{42}$ ; by means of this we can easily determine coefficients of the  $u$ 's and  $v$ 's for Step 3.

*Discussion of Step 3.* The  $c_{ij}$ 's associated with Step 3 are such that each  $\bar{c}_{ij} \geq c_{ij}$ ; thus the procedure ends with this step, and the allocation involved is a solution of the problem. The sum (8.25) obtained is at least as large as the sum associated with any other allocation. From the first column of Step 3 we have that the solution,  $(x_{ij})$ , is as follows:

$$(x_{ij}) = \begin{pmatrix} .10 & 0 & .30 \\ 0 & .20 & 0 \\ .20 & 0 & 0 \\ .05 & .15 & 0 \end{pmatrix}. \quad (3.7)$$

An optimal classification of the group of persons is formed when the persons of type  $i$  assigned to jobs of type  $j$  constitute a proportion  $x_{ij}$  of the total group, where  $x_{ij}$  is given in (3.7) ( $i = 1, \dots, 4; j = 1, \dots, 3$ ). The average productivity associated with this optimal classification is 8.25.

*4. Extensions of the Problem.* Two different extensions will be presented. The first involves a population of personnel categories with which an  $n$ -dimensional distribution is associated; the second involves replacement of the scalar,  $c_{ij}$ , by a vector.

Let  $S$  be an  $n$ -space, and represent any point in  $S$  by  $(z_1, \dots, z_n)$ . Represent the  $i$ th row of  $(c_{ij})$  as a point,  $(c_{i1}, \dots, c_{in})$ , in  $S$  ( $i = 1, \dots, m$ ) and let  $F_m(z_1, \dots, z_n)$  be the distribution function associated with the  $m$  points representing the  $m$  rows of  $(c_{ij})$ ; thus, in the problem associated with (2.2)

$$F_m(z_1, \dots, z_n) = \sum_q a_q, \quad (4.1)$$

where  $q$  ranges over all values of  $i$  such that for every  $j$ ,  $c_{ij} \leq z_j$ .  $F_m(z_1, \dots, z_n)$  is the probability that the  $n$  productivities, say  $c_1, \dots, c_n$ , of an individual selected at random would be such that  $c_j \leq z_j$  ( $j = 1, \dots, n$ ).  $F_m(z_1, \dots, z_n)$  is a discontinuous distribution; an extension of the problem arises when the population is such that the distribution, say  $F(z_1, \dots, z_n)$ , is continuous (e.g., a cumulative normal  $n$ -variate distribution). This extension is considered in (2), (6), and (8). The problem is to find  $n$  mutually ex-

clusive regions, say  $R_1, \dots, R_n$ , in  $S$  such that  $\sum_i \int_{R_i} z_i dF$  is maximized subject to the restrictions:

$$\int_{R_j} dF = b_j, \quad (b_j \geq 0; \sum_j b_j = 1; j = 1, \dots, n). \quad (4.2)$$

The optimal classification of the population would consist in assigning all individuals in  $R_j$  to jobs in the  $j$ th category ( $j = 1, \dots, n$ ). It follows from (6) that the regions would be determined by solving  $n$  simultaneous equations in  $n$  unknowns. A solution can be obtained very easily when  $n = 2$  (e.g., see 8, pp. 218-219). An approximate solution would be obtained by "grouping"—i.e., by approximating  $F(z_1, \dots, z_n)$  by a discontinuous distribution, say  $F_m(z_1, \dots, z_n)$ —and then solving a problem of the type associated with (2.2). The quantity  $a_i$  in (2.2) would represent the amount of a "jump" in the approximating distribution,  $F_m(z_1, \dots, z_n)$  ( $i = 1, \dots, m$ ).

In the first problem stated in section 2, let  $N$  be large and let  $c_{ij}$  be replaced by a vector  $(c_{1,ij}, \dots, c_{k,ij})$  ( $i = 1, \dots, m; j = 1, \dots, n$ ). (See the remarks in the next-to-last paragraph of section 1 regarding a case where  $k$  would equal 2.) Let  $T_h = \sum_{i,j} c_{h,ij} x_{ij}$  ( $h = 1, \dots, k$ ); and let  $R = \sum_h g_h T_h$  be a composite indicator of the value of an allocation  $(x_{ij})$  ( $g_1, \dots, g_k$  are weights). Note that  $R = \sum_{i,j} w_{ij} x_{ij}$ , where  $w_{ij} = \sum_h g_h c_{h,ij}$ ;  $w_{ij}$  is like the  $c_{ij}$  in section 2. The problem is as follows: Find an allocation for which  $R$  assumes its maximum value, where  $x_{11}, \dots, x_{mn}, a_1, \dots, a_m, b_1, \dots, b_n$  are subject to (2.2) and to inequalities of the form  $T_h \geq e_h$  (here  $e_h$  represent a minimum acceptable value of  $T_h$ ; however, the problem is essentially unchanged if for any  $h$  the inequality is reversed). A preliminary is to determine whether there is any allocation satisfying the specified conditions.

The problem stated above is an example of a linear programming problem, which can be solved by means of the simplex method (see 1, Chapters XX, XXI, XXII).

#### REFERENCES

1. Activity analysis of production and allocation (T. C. Koopmans, Editor). Cowles Commission Monograph No. 13. John Wiley and Sons, New York, 1951.
2. Brogden, Hubert E. An approach to the problem of differential prediction. *Psychometrika*, 1946, 11, 139-154.
3. Brown, G. W., and von Neumann, J. Solutions of games by differential equations in Contributions to the theory of games, Annals of Mathematics Study No. 24, edited by H. W. Kuhn and A. W. Tucker. Princeton: Princeton Univ. Press, 1950.
4. Dantzig, George B. Simplex solution of the Hitchcock-Koopmans transportation problem, Headquarters USAF Comptroller, February, 1950 (revised).
5. Hall, Philip. On representatives of subsets. *J. London math. Soc.*, 1935, 10, 26-30.
6. Lord, Frederic M. Notes on a problem of multiple classification. *Psychometrika*, 1952, 17, 297-304.
7. von Neumann, J. and Morgenstern, O. Theory of games and economic behavior, 2nd ed. Princeton: Princeton Univ. Press, 1947.

8. Thorndike, Robert L. The problem of classification of personnel. *Psychometrika*, 1950, 15, 215-235.

*Manuscript received 11/1/51*

*Revised manuscript received 3/9/52*

*Addition at proof reading:* The simplex method provides a general solution of the basic problem stated in section 2; however, it may be awkward when  $m$  or  $n$  is large (see section 3). P. S. Dwyer has proposed some methods including a modification of the simplex method. T. E. Easterfield's paper (A combinatorial algorithm, *J. Lond. Math. Soc.*, 1946, 21, 219-226) gives another interesting method. Procedures for obtaining approximate solutions are given in section III of a forthcoming bulletin by D. F. Votaw, Jr., and J. T. Dailey (*Assignment of Personnel to Jobs*, Human Resources Research Center, Lackland Air Force Base, San Antonio, Texas).



## THE CLOSURE FACTORS RELATED TO OTHER COGNITIVE PROCESSES\*

CAROL PEMBERTON

UNIVERSITY OF DELAWARE

Twenty-five group tests, assembled with certain hypotheses concerning the first and second closure factors in mind, were administered to 154 subjects, mostly graduate students. The intercorrelations were analyzed factorially, yielding eight factors that were rotated to an oblique simple structure. The factors were interpreted as: speed of closure,  $C_1$ ; flexibility of closure,  $C_2$ ; verbal closure,  $C_3$ ; word fluency,  $W$ ; reasoning,  $R$ ; perceptual speed,  $P$ ; the first space factor,  $S_1$ ; and speed of handwriting,  $H$ . Four second-order factors were tentatively described as analytical ability, synthetic ability, speed of perception, and word fluency. Three of the reasoning tests had their highest loadings on  $C_2$  and one on  $C_3$ , which seems to be evidence that flexibility of closure generalizes in the cognitive domain and is associated with analytical ability.

### *I. Purpose of the Study*

Factors named by Thurstone "speed of closure" and "flexibility of closure" have been isolated in several previous studies (1, 2, 17, 19). The purpose of the present study was to ascertain whether abilities on speed of closure and flexibility of closure tests generalize to other domains. The part of the study to be reported in this paper deals with generalization of the closure factors to tasks requiring higher cognitive functions. The generalization of the closure factors to the conative and affective domains will be discussed in a subsequent paper.

The hope of finding meaningful personality variables through studying perception has been expressed by Klein and Schlesinger (9), and by Frenkel-Brunswick (6), among others.

Thurstone (19) writes:

The first closure factor  $C_1$  (speed of closure) seems to facilitate the making of a closure in an unorganized field, the second closure factor  $C_2$  (flexibility of closure) seems to facilitate the retention of a figure in a distracting field. If this . . . interpretation of the two closure factors has any generality beyond the perceptual domain, then one could imagine that the factor  $C_1$  determines the ease with which the subject can unify a complex situation, whereas the second factor determines the ease with which he can keep in mind its essential features against distraction. . . . The first closure factor might be associated with inductive thinking, whereas the second closure factor might be more associated with deductive thinking.

\*This paper summarizes part of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Chicago. The writer is deeply indebted to Dr. L. L. Thurstone for his generous advice and guidance. The complete study is obtainable on microfilm, from the University of Chicago library, Film No. T1279 (price \$2.15). It includes reproductions of the tests used, score distributions, and plots of the oblique  $V$ -matrices.

The present investigation deals with a factor analysis of the intercorrelations of twenty-five tests selected to include longer forms of tests for both closure factors; inductive and deductive reasoning tests; verbal closure tests; and key tests for the first space factor, perceptual speed, word fluency, and speed of handwriting.

## II. Description of the Tests

1. *Figures.* Twenty rows of geometrical figures are presented, with instructions to mark the figures of the six on the right which can be made the same as the one on the left by sliding them around in the plane of the paper.

2. *Cards.* A hundred and twenty diagrams of differently shaped cards are given and the subject asked which cards of the six in a row can be made the same as the model by sliding them around in the plane of the paper.

3. *Gestalt Completion.* This test is similar to Thurstone's adaptation of the Street Gestalt Completion test (12). Two 24-item forms of this test were devised by the writer and the scores for the two combined.

4. *Mutilated Words.* Two 51-item forms similar to Thurstone's test (17) were prepared by the writer and the scores for the two combined.

5. *Concealed Figures.* This 49-row form of the Gottschaldt Figures Test was devised in the Psychometric Laboratory, University of Chicago. A small design at the beginning of each row is followed by four more complex designs. The subject indicates in which of the more complex diagrams the simple figure occurs.

6. *Designs.* In this test of Thurstone's (15), 300 designs are presented, in 40 of which the Greek capital letter "sigma" is concealed. The task is to mark the figures containing the "sigma."

7. *First Letter.* The subject lists all the words he can that begin with the letter "c."

8. *Suffixes.* The subject lists all the words he can that end with the letters "-tion."

9. *Writing Phrase.* The task here is to write the phrase "Now is the time for all good men" as many times as possible within one minute.

10. *X's.* The subject writes as many X's as possible in 30 seconds.

11. *Identical Forms.* The first figure in each of 60 rows of figures is exactly the same as one of the five numbered figures following it. For each row the subject writes the number of the figure which is identical with the model.

12. *Identical Numbers.* The subject is required to mark the number 927 each time it occurs in 18 columns of three-digit numbers.

13. *Number Series.* The subject is to discover the rule underlying each of 22 series of numbers and fill in the blanks.

14. *Letter Series.* This 25-series test is similar to the previous one, except that letters are used.

15. *Figure Classification*. This is a test of Spearman's adapted by Thurstone (14). Two groups of four figures each occur at the beginning of each line, followed by eight more figures. The subject is to discover the rule by which the first two groups were constructed, and then mark those of the remaining eight figures which belong to Group I.

16. *False Premises*. The subject labels as true or false each of 25 syllogisms composed of absurd-sounding premises and conclusions.

17. *Hidden Words*. This test was devised for the present study. A four-letter word appears at the beginning of each of 36 sentences. The subject is to encircle the four-letter word whenever it occurs, either embedded in a longer word, or in two longer words, ignoring spaces and punctuation. This test represents an attempt to produce a verbal test of flexibility of closure. The sentences were made amusing in an effort to strengthen the Gestalt of the complete sentence.

18. *Incomplete Words*. Individual letters are deleted from 50 common English words and the subject is asked to fill in the missing letters.

19. *Anagrams*. Fifty English words are presented, the letters of which, when rearranged, spell another English word. The subject encircles the letter with which the new word begins. The words to be found are well-known ones, the majority being AA or A in Thorndike's (13) word list.

20. *Scrambled Words*. This is a revision of a test constructed by Botzum (2). The letters of 50 common four-letter words are rearranged so as to form nonsense syllables. The subject encircles the first letter of the unscrambled meaningful word.

21. *Copying*. This test was adapted by Thurstone from a test by MacQuarrie (14). The subject is to copy 36 figures in the dotted spaces provided. Each drawing must have the same size, shape, and position as the original.

22. *Four Letter Words*. This test, devised by Bechtoldt (1), and revised by Botzum (2), consists of rows of evenly-spaced typewritten capital letters. In each row there are a few combinations of four letters that spell words, which are to be encircled.

23. *Backward Writing*. In this test, designed by Thurstone (16), 50 words are typed in the normal manner, and below each are four words reversed as in a mirror. The subject underlines the word corresponding to the normally typed one.

24. *Hidden Pictures*. This test of Thurstone's contains items like those found in children's books, in which one can see objects hidden in the lines of a larger picture. Six pictures, in which 28 hidden persons are to be located, constitute the test.

25. *Sentences*. A test similar to this was used by Meili (10). In the present form, devised for this study, the subject is required to write as many sentences as he can containing each of four given words, then as many as he can containing each of five and six given words.

Data on reliability coefficients, means, standard deviations, time limits, and scoring formulas are summarized in Table 1. The reliabilities are corrected split-half coefficients. As the tests are all speed tests, the reliability coefficients based on parallel forms would probably be lower. Distributions were all approximately normal except for the test Backward Writing, where there was a bunching of high scores as a result of a too liberal time limit.

### III. Subjects and Administration of the Tests

The 154 subjects (122 graduate students and 32 undergraduate students) were volunteers, obtained through announcements on bulletin boards at the University of Chicago. There were 94 men and 60 women. All the divisions and professional schools were represented, but the largest group (43 persons) was from the Social Science Division. Ages ranged from 17 to 59 years, with the majority of cases between 19 and 29, and the average 26.4 years.

TABLE 1  
Information Regarding the Tests

Test	Mean	S. D.	<i>r</i>	Time Limits (minutes)		Scoring Formula
				Fore-Ex.	Test	
1. Figures	20.97	9.40	.96	2	3.5	R - W
2. Cards	27.46	12.09	.96	2	4	R - W
3. Gestalt Completion	35.98	6.94	.80	2	3	R
4. Mutilated Words	72.01	14.14	.87	2	5	R
5. Concealed Figures	100.88	28.70	.94	4	10	R - W
6. Designs	26.94	6.28	.94	2	3.5	R
7. First Letter	56.98	13.36	—	2	5.5	Total
8. Suffixes	28.44	9.00	—	1.5	5	Total
9. Writing Phrase	54.05	6.97	—	1	1	Total
10. X's	62.68	8.22	—	.5	.5	Total
11. Identical Forms	44.68	9.11	.98	3	4	R - W/4
12. Identical Numbers	63.74	7.02	.92	3	2.5	R
13. Number Series	11.66	3.53	.82	4	7	R
14. Letter Series	13.55	4.28	.85	4	6	R
15. Figure Classification	66.79	23.22	.94	6	10.5	R - W
16. False Premises	12.33	5.54	.65	3	6	R - W
17. Hidden Words	68.94	15.64	.98	3	5	R
18. Incomplete Words	29.31	9.80	.95	1.5	5	R
19. Anagrams	18.74	7.59	.88	4	6.5	R
20. Scrambled Words	27.21	10.65	.94	4	5	R - W/3
21. Copying	16.36	7.93	.96	4	4	R
22. Four-Letter Words	39.32	9.11	.92	3	4.5	R
23. Backward Writing	40.74	7.98	.98	2	3	R - W/3
24. Hidden Pictures	13.31	3.35	.66	5	14	R
25. Sentences	12.60	3.71	—	1	9	Total

Tests were administered in group form on three consecutive evenings during August, 1950, from 7:30 to 10 p.m.

#### IV. *The Factor Analysis*

All tests were scored by hand twice, usually by two different people. Pearson product-moment correlation coefficients were calculated from the raw scores by I.B.M. equipment (Table 2). Eight factors were extracted from the correlation matrix by the complete centroid method (18). The matrix was factored twice. During the first factoring, the highest value in each column, adjusted after the extraction of each factor, was used as the diagonal entry. During the second factoring, the communalities resulting from the first factoring were employed. The largest difference between the communality estimates resulting from the two factorings is .06, and the majority of differences are .00 or .01. The eighth-factor residuals (Table 3) range from .08 to -.10 with a mean of .00 and a standard deviation of .030. The orthogonal centroid matrix  $F$  is reproduced in Table 4.

The  $F$ -matrix was rotated to a simple structure using both radial and single-plane methods of rotation (18). Table 5 presents the final transformation matrix and Table 6 the resulting oblique factor matrix,  $V$ . Table 7 shows the cosines between the reference vectors of the  $V$ -matrix, and Table 8 the correlations between the primary axes. The reference vectors have been identified by the letters A to H, but in Tables 8-12, which refer to the primary vectors, the columns are denoted by letters which refer to the psychological interpretation of the factors. The columns are kept in the same order to facilitate the comparison of tables.

Table 8 shows that some of the correlations between the primary vectors are high. For this reason the  $A$ -matrix presented in Table 11 was calculated. This table represents the test vectors as linear combinations of the primary vectors. It is obtained by post-multiplying the  $V$ -matrix by the inverse of the diagonal matrix,  $D$ . The values of  $D$  (Table 9) are also used in computing the matrix  $T$  (Table 10), which represents the eight direction cosines of each primary vector. Table 12 gives the correlations between each primary factor and each of the tests.

#### V. *The Second-Order Analysis*

The complete centroid method of factoring was used, and Table 8 was factored three times to stabilize the communalities. Table 13 shows the orthogonal matrix  $F_2$ ; the residuals after five factors had been extracted appear in Table 14. The centroid matrix was rotated to an oblique simple structure. The oblique matrix,  $V_2$ , and the transformation matrix,  $\Lambda_2$ , are shown in Tables 15 and 16, and the cosines between the reference vectors in Table 17. Because of the impossibility of determining the second-order primary vectors with much confidence, the correlations between the primaries were not calculated.

TABLE 2  
Table of Intercorrelations,  $R^*$

Test	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1																									
2	73																								
3	44	49																							
4	25	36	37																						
5	47	63	43	46																					
6	36	40	38	41	54																				
7	15	25	06	25	16	34																			
8	08	13	-03	33	04	22	56																		
9	-09	-05	-08	09	-01	08	20	11																	
10	19	12	-10	08	17	18	23	11	34																
11	48	47	55	32	49	49	23	11	04	09															
12	17	19	18	41	26	44	26	23	19	10	35														
13	31	46	21	44	41	33	20	26	-03	03	27	2													
14	34	46	28	43	36	34	34	25	04	-07	39	41	54												
15	32	48	32	36	54	34	23	11	-10	-02	20	11	45	45											
16	08	19	15	23	36	10	11	04	-06	06	09	-07	40	25	41										
17	30	31	28	41	37	53	42	32	22	17	50	55	24	50	20	12									
18	21	31	13	59	34	33	41	42	12	12	17	39	47	48	33	27	46								
19	21	22	15	46	24	35	51	38	09	17	23	37	33	51	28	26	55	63							
20	25	20	03	44	24	31	52	42	00	12	17	39	45	49	27	21	52	63	73						
21	32	51	38	47	70	45	16	08	13	16	34	26	35	30	40	33	31	39	26	22					
22	20	28	23	49	44	51	48	25	20	24	38	44	25	39	20	12	64	49	52	49	34				
23	36	43	29	45	42	44	45	37	15	19	44	42	33	44	21	16	53	50	45	53	35	40			
24	24	24	39	23	36	24	00	-10	-07	09	27	16	10	08	08	-03	17	09	01	-03	35	21	18		
25	13	23	13	18	18	13	37	20	34	09	28	18	12	27	15	11	29	15	14	07	18	28	17		

\*Decimal points omitted.

# VI. Interpretation of the First-Order Factors

Primary factors are usually interpreted by referring to loadings on the reference vectors. However, the aim of factor analysis is to express the test vectors as linear combinations of the primary vectors, and this is given by the *A*-matrix (Table 11) rather than the *V*-matrix. In an oblique solution the *A*-matrix is proportional to the *V*-matrix by columns but not by rows. A test may have its highest loading on a given reference vector, but its highest component is not necessarily of the corresponding primary vector.

For example, Test 4 had its highest loading on reference vector F, and its second highest on A, whereas in the *A*-matrix its largest component is of the primary vector  $C_3$ , which corresponds to the reference vector A. A more marked reversal occurs in the case of Test 14. Due to our inadequate knowledge of the significance of factor loadings, it is difficult to say whether these reversals are important. However, because it seems more logical to interpret from the *A*-matrix, this has been done. Since most people interpret from the *V*-matrix, the loadings on the reference vectors, as well as the components of the primary vectors, will be quoted.

TABLE 3  
Distribution of Eighth-Factor Residuals\*

Residual	Frequency
.08	4
.07	8
.06	19
.05	18
.04	27
.03	47
.02	80
.01	90
.00	104
-.01	59
-.02	52
-.03	52
-.04	26
-.05	22
-.06	9
-.07	0
-.08	4
-.09	0
-.10	4
N = 625	

\*This table includes the residuals from the diagonal entires.

TABLE 2  
Table of Intercorrelations,  $R^*$

Test	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1		73	44	25	47	36	15	08	-09	19	48	17	31	34	32	08	30	21	21	25	32	20	36	24	13
2	73		49	36	63	40	25	13	-05	12	47	19	46	46	48	19	31	31	22	29	51	28	43	21	23
3		49		37	43	38	06	-03	-08	-10	55	18	21	28	32	15	28	13	15	03	38	23	29	39	13
4	25	36	37		46	41	25	33	09	08	32	41	44	43	36	23	41	59	46	44	47	49	45	23	18
5	47	63	43	46		54	16	04	-01	17	49	26	41	36	54	36	37	34	24	24	70	44	42	36	18
6	36	40	38	41	54		34	22	08	18	49	44	33	34	34	10	53	33	35	31	45	51	44	24	13
7	15	25	06	25	16	34		56	20	23	23	26	20	34	23	11	42	41	51	52	16	48	45	00	37
8	08	13	-03	33	04	22	56		11	11	11	23	26	25	11	04	32	42	38	42	08	25	37	-10	20
9	-09	-05	-08	09	-01	08	20	11		34	04	19	-03	04	-10	-06	22	12	09	00	13	20	15	-07	34
10	19	12	-10	08	17	18	23	11	34		09	10	03	-07	-02	06	17	12	17	12	16	24	19	09	09
11	48	47	55	32	49	49	23	11	04	09		35	27	39	20	09	50	55	39	37	39	26	44	42	16
12	17	19	18	41	26	44	26	23	19	10	35		22	41	11	-07	55	39	37	39	26	44	42	16	18
13	31	46	21	44	41	33	20	26	-03	03	27	2		54	45	45	25	50	48	51	49	30	39	44	08
14	34	46	28	43	36	34	34	25	04	-07	39	41	54		45	25	50	48	51	49	30	39	44	08	27
15	32	48	32	36	54	34	23	11	-10	-02	20	11	45	45		41	12	27	26	21	33	12	16	-03	11
16	08	19	15	23	36	10	11	04	-06	06	09	-07	40	25	41		12	27	26	21	33	12	16	-03	11
17	30	31	28	41	37	53	42	32	22	17	50	55	24	50	20	12		46	55	52	31	64	53	17	29
18	21	31	13	59	34	33	41	42	12	12	17	39	47	48	33	27	46		63	63	39	49	50	09	15
19	21	22	15	46	24	35	51	38	09	17	23	37	33	51	28	26	55	63		73	22	49	53	-03	07
20	25	29	03	44	24	31	52	42	00	12	17	39	45	49	27	21	52	63	73		22	49	53	-03	07
21	32	51	38	47	70	45	16	08	13	16	34	26	35	30	40	33	31	39	26	22		34	40	21	28
22	20	28	23	49	44	51	48	25	20	24	38	44	25	39	20	12	64	49	52	49	34		40	21	28
23	36	43	29	45	42	44	45	37	15	19	41	42	33	44	21	16	53	50	45	53	40		35	21	18
24	24	21	39	23	36	24	00	-10	-07	09	27	16	10	08	08	-03	17	09	01	-03	35		18	17	12
25	13	23	13	18	18	13	37	20	34	09	28	18	12	27	15	11	29	15	14	07	18		17		

\*Decimal points omitted.

### VI. Interpretation of the First-Order Factors

Primary factors are usually interpreted by referring to loadings on the reference vectors. However, the aim of factor analysis is to express the test vectors as linear combinations of the primary vectors, and this is given by the *A*-matrix (Table 11) rather than the *V*-matrix. In an oblique solution the *A*-matrix is proportional to the *V*-matrix by columns but not by rows. A test may have its highest loading on a given reference vector, but its highest component is not necessarily of the corresponding primary vector.

For example, Test 4 had its highest loading on reference vector F, and its second highest on A, whereas in the *A*-matrix its largest component is of the primary vector  $C_3$ , which corresponds to the reference vector A. A more marked reversal occurs in the case of Test 14. Due to our inadequate knowledge of the significance of factor loadings, it is difficult to say whether these reversals are important. However, because it seems more logical to interpret from the *A*-matrix, this has been done. Since most people interpret from the *V*-matrix, the loadings on the reference vectors, as well as the components of the primary vectors, will be quoted.

TABLE 3  
Distribution of Eighth-Factor Residuals\*

Residual	Frequency
.08	4
.07	8
.06	19
.05	18
.04	27
.03	47
.02	80
.01	90
.00	104
-.01	59
-.02	52
-.03	52
-.04	26
-.05	22
-.06	9
-.07	0
-.08	4
-.09	0
-.10	4
N = 625	

\*This table includes the residuals from the diagonal entires.

1. *Verbal Closure, C<sub>3</sub>*

Test	Loadings on A	Components of C <sub>3</sub>
19. Anagrams	.54	.91
20. Scrambled Words	.52	.88
18. Incomplete Words	.48	.81
4. Mutilated Words	.38	.64
22. Four-Letter Words	.36	.61
17. Hidden Words	.34	.57
14. Letter Series	.33	.56
12. Identical Numbers	.32	.54

The tests which are high on this factor all require speed in organizing or reorganizing a set of highly practiced symbols, such as letters or numbers. No spatial or pictorial tests are represented. The factor has a high correlation (.70) with the word fluency factor, W, and the four tests with the highest loadings all require the subject to think of words rapidly, and to select the one which fits the restrictions set by the problem. However, these restrictions are more rigid than those for the W tests. In the tasks high on C<sub>3</sub> only *one* word

TABLE 4  
Orthogonal Centroid Factor Matrix, *F*

Test	I	II	III	IV	V	VI	VII	VIII	<i>h</i> <sup>2</sup>
1	.53	.38	.05	.11	.30	.16	.27	.21	.67
2	.66	.43	.13	-.06	.28	.06	.16	.15	.77
3	.46	.47	-.13	.28	.18	-.17	-.09	.05	.60
4	.68	-.03	.04	.10	-.21	-.27	-.15	.11	.63
5	.70	.48	-.12	-.14	-.23	.07	.10	-.09	.83
6	.65	.10	-.22	.07	-.06	.09	.07	-.20	.54
7	.55	-.41	.06	-.20	.18	-.08	.23	-.12	.62
8	.40	-.42	.21	-.10	.16	-.25	.10	-.21	.53
9	.17	-.37	-.39	-.39	.04	.05	-.19	.20	.55
10	.23	-.13	-.20	-.27	-.07	.18	.31	.18	.35
11	.59	.23	-.31	.18	.30	.06	-.04	-.11	.64
12	.53	-.25	-.25	.26	-.04	.03	-.10	-.11	.50
13	.57	.14	.39	-.02	-.09	.06	-.17	-.06	.54
14	.66	-.04	.25	.14	.15	.15	-.31	-.03	.66
15	.51	.32	.34	-.14	-.07	.08	-.16	-.08	.54
16	.32	.18	.29	-.24	-.20	.04	-.20	-.04	.36
17	.71	-.26	-.24	.21	.04	.19	-.01	-.08	.72
18	.67	-.28	.26	.04	-.22	-.12	-.02	.11	.67
19	.65	-.38	.23	.18	-.14	.07	.05	.07	.68
20	.63	-.39	.37	.20	-.13	.11	.20	-.03	.80
21	.62	.33	-.12	-.19	-.26	-.12	.02	.02	.63
22	.67	-.24	-.21	.07	-.17	.08	.03	.03	.59
23	.68	-.12	-.02	.10	.11	-.06	.12	-.11	.53
24	.27	.28	-.33	.14	-.11	-.27	.05	.09	.38
25	.36	-.11	-.15	-.30	.28	-.04	-.20	.05	.38

will fit the requirements, while in the W tests there are numerous acceptable words.

Mutilated Words and Incomplete Words both have loadings on  $C_1$  (identified as speed of closure), as well as on  $C_3$ . In these two tests the elements to be combined are given in their correct order, and the task can be performed by a rapid synthesis. However, while administering Mutilated Words as an individual test, the writer observed that many subjects did this task analytically, one subject remarking that he was "learning to break the code." Moreover, if the word is not seen quickly and correctly as a whole, then the ability to change one's set rapidly seems to be important. Due to inability to relinquish an incorrect response, many people perform inadequately.

In Anagrams the subject must abandon the Gestalt which is formed by one English word and reorganize the letters into another word. Similarly, in Scrambled Words some of the nonsense syllables are pronounceable and form Gestalts that must be broken. In Four-Letter Words, Hidden Words, and Identical Numbers, one must find or retain a Gestalt in a distracting field. All these tests therefore seem to require elements of what has previously been called "flexibility of closure." The factor itself correlates .54 with flexibility of closure,  $C_2$ .

Letter Series has almost equal components of this factor and factor R. If R accounts for the synthetic aspects of reasoning, then one would expect Letter Series also to have a loading on a factor representing analytic ability. Similarly, Number Series, Figure Classification, and False Premises had loadings on R and  $C_2$ . What differentiates  $C_3$  from  $C_2$  seems to be the more highly practiced nature of the symbols involved in the  $C_3$  tasks.

There is a slight negative correlation between the  $C_3$  factor and  $C_1$ . This may be due to  $C_1$  tests being more pictorial and spatial in nature, or it may be that the  $C_3$  tests require flexibility of closure rather than speed of closure. Most of the tests do seem to require flexibility of closure; but in Botzum's study, Four-Letter Words and Scrambled Words had loadings of

TABLE 5  
Final Transformation Matrix, A

	A	B	C	D	E	F	G	H
I	.266	.170	.215	.113	.166	.123	.112	.097
II	-.389	.359	.001	-.196	.302	.065	.104	-.227
III	.171	.028	-.447	.121	.053	-.076	.336	-.302
IV	.529	-.505	.316	-.397	.044	.167	-.159	-.543
V	-.355	-.477	.061	.430	.535	-.228	.198	.182
VI	.145	.157	.427	-.559	.109	-.809	.219	.103
VII	-.219	-.070	-.208	.180	.413	-.093	-.834	-.396
VIII	.523	-.574	-.652	-.497	.639	.481	-.237	.594

TABLE 6  
Oblique Factor Matrix,  $V = FA$

Test	A	B	C	D	E	F	G	H
1. Figures	.03	-.08	.02	-.07	.64	-.02	-.08	-.02
2. Cards	-.06	.08	-.03	.07	.56	.03	.07	.04
3. Gestalt Completion	.02	-.04	.17	-.04	.30	.27	.07	-.09
4. Mutilated Words	.38	.06	-.01	.02	-.02	.43	.07	.07
5. Concealed Figures	-.07	.53	.21	-.05	.12	.05	.01	-.06
6. Designs	.05	.26	.41	.03	.01	-.05	-.01	-.08
7. First Letter	.02	.00	.03	.45	.07	-.06	-.07	.10
8. Suffixes	.03	-.03	-.02	.53	-.09	.04	.03	-.03
9. Writing Phrase	.06	-.03	.02	.06	-.04	.02	.04	.64
10. X's	.01	.06	-.06	-.03	.20	-.08	-.29	.25
11. Identical Forms	-.04	.01	.44	.05	.24	-.02	.09	.01
12. Identical Numbers	.32	-.05	.41	-.02	.12	.05	.01	.02
13. Number Series	.19	.27	.03	.02	.01	.00	.36	-.06
14. Letter Series	.33	.02	.23	-.01	.07	-.06	.46	.07
15. Figure Classification	.03	.39	.03	.03	.05	-.04	.38	-.04
16. False Premises	.04	.39	-.07	.01	-.10	.00	.33	.06
17. Hidden Words	.34	-.03	.46	-.03	.03	-.08	.01	.07
18. Incomplete Words	.48	.03	-.10	.06	-.03	.25	.05	.05
19. Anagrams	.54	-.08	.06	-.02	.01	.07	.01	-.01
20. Scrambled Words	.52	-.04	.05	.05	.02	-.04	-.04	-.18
21. Copying	.00	.41	.04	.01	.06	.24	-.01	.07
22. Four-Letter Words	.36	.06	.26	-.06	-.02	.08	-.08	.13
23. Backward Writing	.15	.01	.21	.21	.11	.05	-.02	-.05
24. Hidden Pictures	.02	.02	.06	-.05	.11	.39	-.22	-.03
25. Sentences	-.08	.01	.06	.25	.11	.01	.23	.42

.34 and .26 on  $C_1$ , whereas in this study their loadings on  $C_1$  are negligible. Due to this ambiguity the factor has merely been named *Verbal Closure*, although a more accurate descriptive phrase might be "flexibility of closure, utilizing highly practiced symbols."

## 2. Flexibility of Closure, $C_2$

Test	Loadings on B	Components of $C_2$
5. Concealed Figures	.53	.79
21. Copying	.41	.61
15. Figure Classification	.39	.58
16. False Premises	.39	.58
13. Number Series	.27	.40
6. Designs	.26	.39

This factor seems to correspond to Thurstone's flexibility of closure,  $C_2$  (17). Three of the tests are ones which defined this factor in the mechanical aptitude study (19), namely Copying, Designs, and Concealed Figures (a

TABLE 7  
Reference Vector Cosines,  $C = A'A$

	A	B	C	D	E	F	G	H
A	1.000	-.450	-.108	-.616	.028	.318	-.017	.123
B	-.450	1.000	.291	.132	-.518	-.331	.280	-.183
C	-.108	.291	1.000	-.082	-.397	-.541	.257	-.265
D	-.616	.132	-.082	1.000	-.125	.025	.027	-.111
E	.028	-.518	-.397	-.125	1.000	.103	-.304	.232
F	.318	-.331	-.541	.025	.103	1.000	-.291	.127
G	-.017	.280	.257	.027	-.304	-.291	1.000	.221
H	.123	-.183	-.265	-.111	.232	.127	.221	1.000

longer form of Gottschaldt Figures). In addition to these three tests we find three reasoning tests with significant loadings.

The  $C_2$  factor is interesting, since it cuts across the content of the test material. What seems to be required in all these tests is "freedom from *Gestaltbindung*," the Gestalt being formed either by the objective stimulus or by the mental set adopted by the subject. In Concealed Figures one must break the Gestalt formed by the large figure in order to find the small figure. One must also ignore the optical illusions which are often created by the large figure, distorting the appearance of the small figure.

It has been suggested that Copying has a loading on  $C_2$  because the subject must not be influenced by the regularly spaced dots in reproducing the given figure (19). This hypothesis could be investigated by giving the test without the dotted background, as well as with a more highly distracting background. The former method of presentation should decrease its loading on  $C_2$ , the latter increase it. However, it is possible that this test would have a high loading on  $C_2$ , whatever the method of presentation. The  $C_2$  factor seems to represent the ability to abstract. According to Goldstein and

TABLE 8  
Correlations between Primary Factors,  $R = D(A'A)^{-1}D$

	$C_3$	$C_2$	P	W	$S_1$	$C_1$	R	H
$C_3$	1.000	.538	.078	.704	.348	-.269	-.182	.052
$C_2$	.538	1.000	.117	.333	.540	.092	-.175	.095
P	.078	.117	1.000	.169	.353	.465	-.099	.204
W	.704	.333	.169	1.000	.304	-.180	-.155	.120
$S_1$	.348	.540	.353	.304	1.000	.190	.158	-.134
$C_1$	-.269	.092	.465	-.180	.190	1.000	.241	-.068
R	-.182	-.175	-.099	-.155	.158	.241	1.000	-.376
H	.052	.095	.204	.120	-.134	-.068	-.376	1.000

TABLE 9  
Diagonal Entries of  $D_{1p}$ 

	$C_3$	$C_2$	P	W	$S_1$	$C_1$	R	H
$d$	.594	.673	.740	.688	.706	.743	.834	.877

Scheerer (7) a characteristic of abstract behavior is "to grasp the essential of a given whole; to break up a given whole into parts, to isolate and to synthesize them." This seems a very good description of what one must do in Copying.

In Figure Classification the subject must not be bound by his first hypothesis concerning the classification of the figures into two groups, nor persevere with previous solutions inapplicable to subsequent items. In Number Series the subject must also test one hypothesis after another, applying the four arithmetical operations of addition, subtraction, multiplication, and division. In False Premises the subject must not be influenced by the surface absurdity of the statements. "A monkey hanging from the branch of a tree is a fruit," sounds a ridiculous assertion, but in terms of the given premises the conclusion is to be marked "True."

The Designs Test has a relatively low loading on this factor, its highest loading being on perceptual speed, P. Designs was devised as a perceptual speed test but in the first study in which it was used (15) it did not have a loading on P. Thurstone later hypothesized that Designs would have loading on  $C_2$ , and this was borne out in the mechanical aptitude study (19). However, it now appears that for a more highly educated group of subjects the test is one which requires perceptual speed to a greater extent than flexibility of closure.

The two reasoning tests which have their highest loadings on  $C_2$  are the two tests which had the highest loadings on Botzum's deduction factor. This tends to corroborate Thurstone's hypothesis that  $C_2$  is associated with deduc-

TABLE 10  
Direction Cosines of Primary Vectors,  $T = DA^{-1}$ 

	I	II	III	IV	V	VI	VII	VIII
$C_3$	.741	-.480	.350	.078	-.219	.147	.131	.074
$C_2$	.711	.279	.112	-.386	-.452	.110	.187	-.060
P	.552	.092	-.678	.331	.202	.033	-.124	-.248
W	.647	-.504	.229	-.194	.256	-.249	.219	-.249
$S_1$	.717	.473	.129	.037	.339	.141	.259	.207
$C_1$	.243	.494	-.354	.227	-.098	-.634	-.312	.109
R	-.010	.380	.492	.122	.334	-.026	-.695	-.051
H	.144	-.386	-.581	-.579	.031	.127	-.201	.316

tive reasoning (19). The writer prefers to differentiate the reasoning factors in terms of the methods of problem-solving involved, rather than in the epistemological terms of induction and deduction. Yela (20) maintains that the solution to reasoning problems can occur either "by closure of the unfinished configuration, perceiving the lacking element as connected with the others, or by analytically reasoning out the principle connecting the items." This is not a new idea. Duncker (5) differentiates "analytic and synthetic evidence," which, he contends, are not opposites. Synthetic evidence can occur because "not all possible aspects of a thought object are necessary to its construction, just as little as all possible aspects of a visual object are necessary to the unambiguous comprehension of its structure." Duncker maintains, however, that the inspectional (synthetic) technique can become a handicap in problem solving, as the thought material is then more thoroughly imbued with perceptual functions, which may be a hindrance in restructuring problems. This observation is interesting in the light of the slight negative correlation of both  $C_2$  and  $C_3$  with synthetic reasoning, R.

TABLE 11  
Test Vectors as Linear Combinations of Primary Vectors,  $A = VD^{-1}$

	$C_3$	$C_2$	P	W	$S_1$	$C_1$	R	H
1. Figures	.05	-.12	.03	-.10	.91	-.03	-.10	-.02
2. Cards	-.10	.12	-.04	.10	.79	.04	.08	.05
3. Gestalt Completion	.03	-.06	.23	-.06	.43	.36	.08	-.10
4. Mutilated Words	.64	.09	-.01	.03	-.03	.58	.08	.08
5. Concealed Figures	.12	.79	.28	-.07	.17	.07	.01	-.07
6. Designs	.08	.39	.55	.04	.01	-.07	-.01	-.09
7. First Letter	.03	.00	.04	.65	.10	-.08	-.08	.11
8. Suffixes	.05	-.04	-.03	.77	-.13	.05	.04	-.03
9. Writing Phrase	.10	-.04	.03	.09	-.06	.03	.05	.73
10. X's	.02	.09	-.08	-.04	.28	-.11	-.35	.29
11. Identical Forms	-.07	.01	.59	.07	.34	-.03	.11	.01
12. Identical Numbers	.54	-.07	.55	-.03	-.17	.07	.01	.02
13. Number Series	.32	.40	.04	.03	.01	.00	.43	-.07
14. Letter Series	.56	.03	.31	-.01	.10	-.08	.55	.08
15. Figure Classification	.05	.58	.04	.04	.07	-.05	.46	-.05
16. False Premises	.07	.58	-.09	.01	-.14	.00	.40	.07
17. Hidden Words	.57	-.04	.62	-.04	.04	-.11	.01	.08
18. Incomplete Words	.81	.04	-.14	.09	-.04	.34	.06	.06
19. Anagrams	.91	-.12	.08	-.03	.01	.09	.01	-.01
20. Scrambled Words	.88	-.06	.07	.07	.03	-.05	-.05	-.21
21. Copying	.00	.61	.05	.01	.09	.32	-.01	.08
22. Four-Letter Words	.61	.09	.35	-.09	-.03	.11	-.10	.15
23. Backward Writing	.25	.01	.28	.31	.16	.07	-.02	-.06
24. Hidden Pictures	.03	.03	.08	-.07	.16	.52	-.26	-.03
25. Sentences	-.13	.01	.08	.36	.16	.01	.28	.48

Burt (3), in describing two types of problem-solvers, writes: "... the characteristic procedure of the former group is explicit or analytic reasoning, depending on a succession of logical steps; of the latter, an implicit or synthetic apperception of what later psychologists have called Gestalten or configurations, depending on an intuitive insight which embraces the component aspects almost simultaneously."

It would seem that the  $C_2$  factor is associated with the more analytical method of problem-solving, whereas factor R may represent the synthetic ability discussed by Burt, Duncker, and Yela, among others.

The  $C_2$  factor seems to correspond closely with Meili's plasticity factor, where the Gottschaldt Figures, Kohs' Blocks, problems taken from Duncker, and Number Series all have significant loadings (10). It is also similar to Factor A in Rimoldi's study (11).

### 3. *Perceptual Speed, P*

Test	Loadings on C	Components of P
17. Hidden Words	.46	.62
11. Identical Forms	.44	.59
12. Identical Numbers	.41	.55
6. Designs	.41	.55
22. Four-Letter Words	.26	.35

This factor corresponds with Thurstone's perceptual speed factor, P (14, 15). The test with the highest loading on this factor is Hidden Words. This test was constructed in an effort to obtain a verbal equivalent of the Concealed Figures test. At the time of construction the writer hypothesized that the test might have its chief loading on perceptual speed, since it is difficult to write sentences with such strong Gestalt properties that one's perception of the embedded word is hindered. It was, in fact, gratifying to find that Hidden Words has almost equal components of  $C_3$  and P.

The tests high on this factor require the subjects to find a given stimulus among distracters, which do not have strong Gestalt properties that distort or conceal the desired stimulus.

### 4. *Word Fluency, W*

Test	Loadings on D	Components of W
8. Suffixes	.53	.77
7. First Letter	.45	.65
25. Sentences	.25	.36

The three tests with loadings on this factor define one of the most clear-cut hyperplanes in the simple structure. The tests Suffixes and First Letter were introduced into the present battery as key tests for the word fluency factor, W. Sentences has an obvious "fluency-with-words" component, although it does not require the subject to think of isolated words fitting some restriction, which is how word fluency has been interpreted previously (14).

The test Sentences was introduced because Meili (10), using a similar test and a population of children, had obtained a loading on the globalization factor. In our case the highest loading for this test was on speed of handwriting, H. The test did have a small loading on R, which seems similar to Meili's globalization factor. For adults, well versed in using words, the ability to induce something common to the given words does not seem to be as important as in the case of children.

5. *Space, S<sub>1</sub>*

Tests	Loadings on E	Components of S
1. Figures	.64	.91
2. Cards	.56	.79
3. Gestalt Completion	.30	.43
11. Identical Forms	.24	.34
10. X's	.20	.28

The highest loadings on this factor are the two key tests for the first space factor, *S<sub>1</sub>*. Thurstone interprets this factor as "the ability to visualize

TABLE 12  
Correlations between the Tests and Primary Factors,  $R_{jp} = FT''$

	C <sub>3</sub>	C <sub>2</sub>	P	W	S <sub>1</sub>	C <sub>1</sub>	R	H
1. Figures	.24	.37	.31	.18	.81	.13	.07	-.12
2. Cards	.30	.53	.30	.31	.87	.21	.19	-.08
3. Gestalt Completion	.02	.22	.51	.03	.57	.58	.27	-.18
4. Mutilated Words	.53	.47	.32	.40	.36	.41	.05	.05
5. Concealed Figures	.30	.81	.43	.18	.65	.36	-.04	.02
6. Designs	.38	.50	.57	.33	.46	.21	-.13	.07
7. First Letter	.58	.32	.19	.76	.28	-.19	-.25	.23
8. Suffixes	.49	.16	.07	.72	.11	-.12	-.07	.04
9. Writing Phrase	.13	.06	.18	.19	-.10	-.05	-.25	.73
10. X's	.24	.31	.07	.18	.19	-.18	-.44	.37
11. Identical Forms	.16	.25	.71	.24	.58	.34	.08	.06
12. Identical Numbers	.44	.19	.56	.35	.18	.16	-.13	.16
13. Number Series	.49	.52	.08	.33	.45	.08	.33	-.17
14. Letter Series	.55	.32	.32	.42	.48	.07	.38	-.04
15. Figure Classification	.33	.56	.06	.22	.48	.13	.36	-.15
16. False Premises	.25	.46	-.09	.13	.22	.06	.25	-.03
17. Hidden Words	.60	.33	.64	.48	.38	.05	-.18	.22
18. Incomplete Words	.76	.49	.11	.57	.31	.06	-.04	.05
19. Anagrams	.81	.39	.18	.57	.31	-.11	-.10	.02
20. Scrambled Words	.87	.41	.09	.64	.34	-.25	-.13	-.13
21. Copying	.29	.70	.33	.21	.48	.41	-.06	.12
22. Four-Letter Words	.60	.45	.47	.43	.31	.09	-.28	.28
23. Backward Writing	.54	.38	.44	.57	.47	.11	-.09	.04
24. Hidden Pictures	-.04	.20	.38	-.04	.24	.53	-.11	.02
25. Sentences	.15	.15	.26	.34	.22	.08	.08	.42

a rigid configuration when it is moved into different positions" (19). This ability would certainly be called for by the tests Figures and Cards. In the case of Identical Forms and Gestalt Completion no movement from one position to another occurs. However, Identical Forms has consistently had a small loading on  $S_1$  (1, 2, 14). The loading of Gestalt Completion on this factor has not occurred previously. It may be that in devising new forms of this test the factorial composition has been somewhat changed.

The small loading of X's on  $S_1$  may not have occurred by chance. Chapman (4) in analyzing the MacQuarrie test found a correlation of .68 between the factors he identifies as spatial and motor. The highest loading on the spatial factor was given by the test Dotting, which requires skill similar to that used in making X's rapidly.

6. *Speed of Closure,  $C_1$*

Tests	Loadings on F	Components of $C_1$
4. Mutilated Words	.43	.58
24. Hidden Pictures	.39	.52
3. Gestalt Completion	.27	.36
18. Incomplete Words	.25	.34
21. Copying	.24	.32

This is evidently the factor identified as speed of closure in previous studies (2, 19). Mutilated Words and Gestalt Completion were the two tests which defined this factor before, and Hidden Pictures and Incomplete Words had loadings on it. This factor is represented by tests in which the subject must unify a relatively unstructured perceptual field.

Scrambled Words and Four-Letter Words, which had small loadings on this factor in Botzum's study, do not appear on our  $C_1$  factor. The difference between these two tests and Mutilated Words and Incomplete Words seems to be that in the latter no rearrangement of parts must take place, as in Scrambled Words, or extraction from a background, as in Four-Letter Words.

TABLE 13  
Centroid Factor Matrix,  $F_2$

	I	II	III	IV	V	$h^2$
$C_2$	.80	.15	-.43	.01	-.15	.87
$C_2$	.65	.28	.17	-.36	-.31	.76
P	.30	.31	.50	.40	.14	.62
W	.69	.12	-.34	.23	.08	.67
$S_1$	.45	.61	.09	-.27	.24	.71
$C_1$	-.17	.54	.50	.23	-.06	.63
R	-.41	.50	-.20	-.06	.11	.47
H	.32	-.40	.31	.17	.05	.39

TABLE 14  
Fifth-Factor Residuals in Second Order

	C <sub>3</sub>	C <sub>2</sub>	P	W	S <sub>1</sub>	C <sub>1</sub>	R	H
C <sub>3</sub>		.00	.03	-.01	-.02	.00	.01	-.01
C <sub>2</sub>	.00		-.06	.01	.03	.02	-.01	.03
P	.03	-.06		-.01	.06	.02	-.03	-.02
W	-.01	.01	-.01		-.01	.00	.00	.02
S <sub>1</sub>	-.02	.03	.06	-.01		-.03	.01	-.02
C <sub>1</sub>	.00	.02	.02	.00	-.03		.02	.00
R	.01	-.01	-.03	.00	.01	.02		.01
H	-.01	.03	-.02	.02	-.02	.00	.01	

One might think that Hidden Pictures would require flexibility of closure rather than speed of closure, but this test has consistently had loadings on C<sub>1</sub>. It seems that it is not breaking the Gestalt formed by the large picture, but the sudden combination of the correct elements, which accounts for individual differences in Hidden Pictures. In Concealed Figures and Designs the subject is shown the figure which he is to locate and can therefore perform the task analytically, whereas in Hidden Pictures he merely knows that he is looking for pictures of people. Reproductions of the faces to be found could be given preceding each large picture, to find out whether the test would then have a loading on C<sub>2</sub>.

#### 7. Reasoning, R

Tests	Loadings on C <sub>1</sub>	Components of R
14. Letter Series	.46	.55
15. Figure Classification	.38	.46
13. Number Series	.36	.43
16. False Premises	.33	.40
25. Sentences	.23	.28
24. Hidden Pictures	-.22	-.26
10. X's	-.29	-.35

The three tests with highest loadings on this factor are all classic tests of induction. As previously stated the writer prefers to differentiate reasoning factors in terms of the mental processes involved, rather than in terms of the types of problems involved. Figure Classification, Number Series, and False Premises each have higher components of the C<sub>2</sub> factor, and Letter Series a higher component of C<sub>3</sub>, than they have of R. The abilities represented by C<sub>2</sub> and C<sub>3</sub> were interpreted as being analytical in nature. It may be that the synthetic apperception of the solution may be represented in the present factor. Letter Series would appear to be the test which lends itself most easily to this form of solution. Some subjects perform this test by singing the series sub-vocally, which is a method of solution that depends on

TABLE 15  
Oblique Factor Matrix,  $V_2$ 

	AA	BB	CC	DD	EE*
$C_3$	.15	.03	-.05	.57	-.13
$C_2$	.65	.00	.02	-.02	-.21
P	-.01	-.07	.61	.15	.09
W	-.05	-.01	.07	.59	.07
$S_1$	.55	.37	.02	.01	.25
$C_1$	.03	.30	.62	-.04	-.13
R	-.04	.65	.03	.03	.00
H	-.02	-.59	.09	-.05	.13

\*Residual

perceiving the series as a Gestalt, rather than analyzing the principle upon which it was built. It has already been mentioned that some of the variance of Sentences is accounted for by the ability to induce the common elements which permit the given words to be combined into a meaningful sentence.

Our interpretation of this factor as reasoning "by closure of the unfinished configuration," to use Yela's words (20), seems to be justified by the inter-correlations of this factor with the others, since its highest positive correlation is with speed of closure (.24). This gives some corroborative evidence for

TABLE 16  
Transformation Matrix,  $A_2$ 

	AA	BB	CC	DD	EE
I	.34	-.33	.02	.33	.09
II	.22	.86	.41	.23	-.12
III	.33	-.39	.43	-.56	.08
IV	-.86	-.10	.75	.70	-.14
V	.00	.00	-.30	-.17	.98

TABLE 17  
Reference Vector Cosines,  $C_2$ 

	AA	BB	CC	DD	EE
AA	1.01	.03	-.41	-.62	.15
BB	.03	1.01	.10	.24	-.15
CC	-.41	.10	1.01	.44	-.41
DD	-.62	.24	.44	.99	-.31
EE	.15	-.15	-.41	-.31	1.01

Thurstone's supposition that "the first closure factor might be associated with inductive thinking," although the association is not as strong as the association of the second closure factor with analytical thinking.

This is the only factor in which there are negative loadings to account for. X's was the last test in the battery, and the subjects were aware of this. The brighter subjects may have been bored by this simple task, whereas the less apt subjects may have felt this was "positively their last chance" to make a good score. The negative loading of Hidden Pictures may also be due to motivational factors.

Factor R is similar to Factor G in Rimoldi's study (11), on which Figure Classification, Letter Series, and Number Series all had loadings. Referring to the tests with loadings on G, Rimoldi writes: "... In the easy items the analytical activity is limited and the grouping of the letters and numbers is rather obvious. As soon as the problems increase in difficulty the grouping becomes less obvious, and the subject has to discover the structure characterizing each item." It is probably the easier items which account for the loadings of these tests on R, but further research is needed to verify this.

#### 8. *Speed of Handwriting, H*

Tests	Loadings on H (Reference Vector)	Components of H (Factor)
9. Writing Phrase	.64	.73
25. Sentences	.42	.48
10. X's	.25	.29

This seems to be the same factor as Bechtoldt's (1) factor H, which he identifies as "the ability to operate hand-finger neuromuscular response mechanisms at a high rate of speed for short periods of time." As we only have tests involving handwriting represented here, we have identified the factor by the simpler term "speed of handwriting." It was gratifying to find that this factor had only one positive correlation of any size (.20), this being with perceptual speed. Individual differences in ability to write quickly therefore did not appear to affect the scores of most of the paper-and-pencil tests used in this battery.

#### VII. *Interpretation of the Second-Order Factors*

With only eight primary factors it is not possible to determine the second-order factors with much confidence. The interpretation of these factors is therefore highly tentative.

##### *Factor AA*

Flexibility of Closure	.65
Space	.55

It must be remembered that our flexibility of closure factor had three reasoning tests with high loadings on it. Bearing this in mind, it becomes apparent that this factor is very similar to the second-order factor  $\alpha$ , isolated by Botzum, on which space, deduction, induction, and flexibility of closure all had loadings over .60.

In the mechanical aptitude study, although a second-order analysis was not performed, it seems likely that inductive reasoning, space, and flexibility of closure would come up on one second-order factor, since I and  $C_2$  correlate .63; I and  $S_1$ , .38; and  $S_1$  and  $C_2$ , .53.

The space primary has loadings on both this factor and factor BB. This also occurred in Botzum's study. He points out that one can perform the space tests in two ways. One can either analytically reason out which way the figure would be facing, when turned in the plane of the paper; or one can actually imagine the rotation of the figure. It is our hypothesis that the analytic procedure in problem-solving is what is represented in Factor AA.

*Factor BB*

Reasoning	.65
Space	.37
Speed of Closure	.30
Speed of Handwriting	-.59

This factor seems to represent the ability to solve problems synthetically. We tentatively interpreted our reasoning factor R as the ability to synthesize, and above we mentioned the possibility of solving the space tests by two different methods, one analytic, the other synthetic. The synthetic element in speed of closure is obvious. Both Jones (8) and Rimoldi (11) have reported second-order factors which they interpret as synthetic ability.

The negative loading of speed of handwriting on this factor is difficult to account for. In Bechtoldt's study speed of handwriting correlated -.22 with Factor G, which seems to correspond to our  $C_1$  factor. Future studies will establish whether this negative relationship occurs consistently.

*Factor CC*

Speed of Closure	.62
Perceptual Speed	.61

This factor seems to be speed of perception. It is similar to factor F found in the factorial study of perception (17). It is not a more generalized speed factor, since speed of handwriting does not have a loading on it. Similarly, Thurstone's F was clearly differentiated from reaction time. Dispersed attention seems to be required in both the speed of closure and perceptual speed tests.

*Factor DD*

Word Fluency	.59
Verbal Closure	.57

This factor is interpreted as the ability to think of words rapidly which fit certain formal requirements, such as beginning with a certain letter, ending with a certain suffix, containing certain given letters, etc. If our interpretation of the verbal closure factor was correct, one would expect that  $C_3$  would have a loading on the second-order factor AA as well as DD. However, its loading on AA is only .15. Due to the indeterminacy of the second-order structure, this does not necessarily invalidate our interpretation of  $C_3$ . Furthermore, the cosine of the angle between AA and DD is  $-.62$ ; therefore a positive correlation between the second-order primaries corresponding to AA and DD could be predicted. This is probably due to the flexibility required by all the tests with loadings on these two second-order factors.

Factor EE has been left without interpretation as it appears to be a residual factor.

### VIII. Discussion

The aim of this study was to ascertain whether abilities on speed of closure and flexibility of closure tests would generalize to tasks requiring higher cognitive functions. It yielded strong evidence that flexibility of closure is associated with analytical reasoning. The association is so close that only one factor appears on which all the flexibility of closure tests and three of the reasoning tests have loadings. Correlated .54 with this factor is one in which rearrangement of letters or numbers to form new Gestalts, or the picking out of Gestalts from a distracting background, is required in all the tests. This indicates that there is a connection between perceptual flexibility, the flexibility required to solve analytical reasoning problems, and the flexibility needed to solve problems utilizing highly practiced symbols but where meaning is not important.

The evidence for the generalization of speed of closure in the higher cognitive domain is not so clear-cut as for flexibility of closure. It is only in the second order that we find a coming together of the speed of closure and reasoning tests, but this second-order factor does seem to indicate that both perceptual and reasoning problems can be solved by a rapid synthetic process.

### REFERENCES

1. Bechtoldt, H. P. Factorial study of perceptual speed. Unpublished Ph.D. dissertation, Department of Psychology, University of Chicago, 1947.
2. Botzum, W. A. A factorial study of the reasoning and closure factors. *Psychometrika*, 1951, 16, 361-386.
3. Burt, Cyril. The structure of the mind: a review of the results of factor analysis, II. *Brit. J. educ. Psychol.*, 1949, 19, 176-199.
4. Chapman, R. L. The MacQuarrie test of mechanical ability. *Psychometrika*, 1948, 13, 175-179.
5. Duncker, Karl. On problem-solving. *Psychol. Monogr.*, 1945, 58, No. 5.

6. Frenkel-Brunswik, Else. Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Personal.*, 1949, 18, 108-143.
7. Goldstein, K., and Scheerer, M. Abstract and concrete behavior. *Psychol. Monogr.*, 1941, 53, No. 2.
8. Jones, L. V. Primary mental abilities in the Stanford-Binet, Age 13. Psychometric Laboratory Report, No. 71, University of Chicago, 1951.
9. Klein, G. S., and Schlesinger, H. Where is the perceiver in perceptual theory? *J. Personal.*, 1949, 18, 32-47.
10. Meili, R. L'analyse de l'intelligence. *Archives de Psychologie*, 1946, 31, 2-64.
11. Rimoldi, H. J. A. The central intellective factor. *Psychometrika*, 1951, 16, 75-93.
12. Street, R. T. A Gestalt completion test. Contributions to Education, No. 481, Teachers College, Columbia University, New York, 1931.
13. Thorndike, E. L. A teacher's word book of the twenty thousand words found most frequently and widely in general reading for children and young people. New York: Teachers College, Columbia University. Revised 1932.
14. Thurstone, L. L. Primary mental abilities. *Psychometric Monogr.*, No. 1, Chicago: Univ. Chicago Press, 1938.
15. Thurstone, L. L. The perceptual factor. *Psychometrika*, 1938, 3, 1-17.
16. Thurstone, L. L., and Thurstone, Thelma G. Factorial studies of intelligence. *Psychometric Monogr.*, No. 2, Chicago: Univ. Chicago Press, 1941.
17. Thurstone, L. L. A factorial study of perception. Chicago: Univ. Chicago Press, 1944.
18. Thurstone, L. L. Multiple-factor analysis. Chicago: Univ. Chicago Press, 1947.
19. Thurstone, L. L. Mechanical aptitude III. Analysis of group tests. Psychometric Laboratory Report, No. 55, University of Chicago, 1949.
20. Yela, M. Application of the concept of simple structure to Alexander's data. *Psychometrika*, 1949, 14, 121-135.

*Manuscript received 1/7/52*

*Revised manuscript received 2/21/52*

## SOME EXAMPLES OF MULTIVARIATE ANALYSIS IN EDUCATIONAL AND PSYCHOLOGICAL RESEARCH

FRED. T. TYLER

UNIVERSITY OF CALIFORNIA

This paper provides the formulas necessary for testing the significance of the differences between mean values of different multivariate normal populations by Hotelling's generalization of "Student's" ratio. It also indicates the methods, proposed by Mahalanobis, and by Rao, of classifying different multivariate populations and individuals. The methods are illustrated by means of personality test data obtained from students preparing for the General Elementary and the General Secondary Teaching Credentials.

The multiple regression equation serves two primary purposes: It provides a method for estimating the value of a battery of tests in predicting some criterion, and it furnishes a basis for determining the most appropriate weights to give each subtest in the battery. Problems which employ multiple regression equations, then, are concerned with the prediction of one variable from a series of other variables, all variables being expressed in quantitative terms. Analogous to the problem of prediction is that of classification of individuals where the criterion is more qualitative than numerical in nature. Thus, we may be interested in differentiating between two or more groups of individuals rather than in predicting the numerical value of the dependent variable for each individual. Fisher (1) proposed the discriminant function as a suitable technique for this purpose. The discriminant function provides material for answering the question: "What is the best system of weights to apply to the various independent variables in order to produce a maximum separation between the groups?" As Garrett (4) points out, this becomes a multiple regression problem if the numerical value of the dependent variables for each individual is known.

There is another more general type of question that might be asked: (1) "Could these two groups have arisen from the same population?" This is clearly a question of the allocation of groups to a population rather than of prediction. The problem propounded is not concerned with maximizing the difference between groups, but is simply concerned with the similarity of the two groups; nor is the question related to the prediction of the numerical value of a dependent variable. Having obtained a negative answer to this particular question, the investigator may proceed to ask two other questions: (2) "Can a given individual be classified into either of the two groups under study?" (3) "If so, into which group is he most properly classified?"

Hotelling (6) has proposed a method of answering the original question: Are these two groups significantly different from one another? Mahalanobis (9) and Rao (11, 12, 13) have suggested procedures for answering the other two questions. This paper will describe and illustrate the methods. A complete explanation is not possible within the limits of the paper; certain features of the development must be assumed. The mathematical development of these tests is available in various sources. Analogy of the technique with the  $t$ -ratio for small samples will be pointed out.

*Problem I. To Test the Significance of the Difference Between the Mean Values of Multivariate Normal Populations.*

Typically this problem has been attacked by determining the significance of the difference between the means of the two groups on each variable successively. Such a procedure does not make maximum use of the available data since it does not consider the inter-correlations of the variables in each of the groups being studied. The  $t$ -test considers the variance in each of the variables but does not make use of the covariances of the variables. Hotelling (6) proposed the  $T$ -test to study the significance of the difference between two multivariate normal populations where it can be assumed that the two populations have the same dispersion matrix, i.e., have equal variances and covariances.

Suppose we give the same  $r$  tests to two groups, one group having  $N_x$  subjects and the other  $N_s$ .

Let  $X_i$  ( $i = 1, 2, \dots, r$ ) be the random variable defined by the  $i$ th test in the first group,

$Z_i$  ( $i = 1, 2, \dots, r$ ) be the random variable defined by the  $i$ th test in the second group,

$X_{ik}$  be the score of the  $k$ th individual in the first group on the  $i$ th test,

$Z_{ik}$  be the score of the  $k$ th individual in the second group on the  $i$ th test,

$\bar{X}_i$  be the sample mean of the  $i$ th test in the first group, and

$\bar{Z}_i$  be the sample mean of the  $i$ th test in the second group.

Assume that the dispersion matrices for these tests from the two samples are equal. Then the best (unbiased) estimates of the sample variance for the  $i$ th test is

$$\begin{aligned} a_{ii} &= \frac{1}{n} \left[ \sum_{k=1}^{N_x} (X_{ik} - \bar{X}_i)^2 + \sum_{k=1}^{N_s} (Z_{ik} - \bar{Z}_i)^2 \right] \\ &= \frac{1}{n} \left[ \sum_{k=1}^{N_x} X_{ik}^2 - N_x \bar{X}_i^2 + \sum_{k=1}^{N_s} Z_{ik}^2 - N_s \bar{Z}_i^2 \right], \end{aligned}$$

where  $n = N_x + N_s - 2$ , and  $i = 1, 2, \dots, r$ .

Similarly, the unbiased estimate of the covariance,  $i \neq j$ , becomes

$$\begin{aligned} a_{ij} = a_{ji} &= \frac{1}{n} \left[ \sum_{k=1}^{N_x} (X_{ik} - \bar{X}_i)(X_{jk} - \bar{X}_j) + \sum_{k=1}^{N_z} (Z_{ik} - \bar{Z}_i)(Z_{jk} - \bar{Z}_j) \right] \\ &= \frac{1}{n} \left[ \sum_{k=1}^{N_x} X_{ik}X_{jk} - N_x \bar{X}_i \bar{X}_j + \sum_{k=1}^{N_z} Z_{ik}Z_{jk} - N_z \bar{Z}_i \bar{Z}_j \right], \end{aligned}$$

where  $n = N_x + N_z - 2$ , and  $i, j = 1, 2, \dots, r, i \neq j$ .

Let

$$d_i = \frac{\bar{X}_i - \bar{Z}_i}{\sqrt{\frac{1}{N_x} + \frac{1}{N_z}}} = (\bar{X}_i - \bar{Z}_i) \sqrt{\frac{N_x N_z}{N_x + N_z}}.$$

The dispersion matrix has the determinant:

$$a = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{r1} & a_{r2} & \cdots & a_{rr} \end{vmatrix}.$$

Define  $A_{ij}$  as the co-factor of  $a_{ij}$  in the determinant  $a$ .

The traditional procedure for testing the significance of the difference between two means in educational and psychological research has been by means of the  $t$ -test for single tests,

$$\begin{aligned} t_i &= \frac{\bar{X}_i - \bar{Z}_i}{\sqrt{a_{ii}}} \sqrt{\frac{N_x N_z}{N_x + N_z}} \\ &= \frac{d_i}{\sqrt{a_{ii}}}, \quad (i = 1, 2, \dots, r) \end{aligned}$$

and with  $N_x + N_z - 2$  degrees of freedom.

Hotelling's  $T$ -test for testing the significance of the difference between two multivariate normal populations is

$$T^2 = \frac{1}{a} \left[ \sum_{i=1}^r \sum_{j=1}^r A_{ij} d_i d_j \right],$$

which reduces to  $t_i$  when we consider only the  $i$ th test.

In order to determine the significance of the computed value of  $T^2$ , it is necessary to compute

$$x = \frac{1}{1 + \frac{T^2}{n}}.$$

The probability that a value as large as or larger than  $T^2$  could occur by chance can be determined from Pearson's tables of the Incomplete Beta Function,  $I_x(p, q)$ , where  $p = (n - r + 1)/2$  and  $q = r/2$  (10). The table indicates the probability that the obtained value of  $T$  will be exceeded by chance if in fact there is only one population.

The significance of  $T^2$  may also be determined from the  $F$  tables,

$$F = \left( \frac{n - r + 1}{nr} \right) T^2,$$

with  $r$  degrees of freedom for  $n_1$  (the larger variance) and  $(n - r + 1)$  degrees of freedom for  $n_2$  (the smaller variance).

*Problem II. The Classification of an Individual into Either of Two Populations.*

If we reject the hypothesis that two samples are members of a single population, it sometimes happens that we wish to know whether a certain individual can be classified into either of the two specified populations. The procedure available for such a problem is an extension of Problem I to consider the case where three groups are under consideration, one group consisting of the individual to be classified (13). As in the preceding problem, it is necessary to set up the dispersion matrix of the unbiased estimate of the variances and covariances.

Let  $Y_i$  ( $i = 1, 2, 3, \dots, r$ ) be the random variable defined by the  $i$ th test in the third group,

$Y_{ik}$  be the score of the  $k$ th individual in the third group on the  $i$ th test, and

$\bar{Y}_i$  be the sample mean of the  $i$ th test in the third group.

Then the unbiased estimate of the sample variance for the  $i$ th test is

$$a_{ii} = \frac{1}{n} \left[ \sum_{k=1}^{N_x} (X_{ik} - \bar{X}_i)^2 + \sum_{k=1}^{N_z} (Z_{ik} - \bar{Z}_i)^2 + \sum_{k=1}^{N_y} (Y_{ik} - \bar{Y}_i)^2 \right],$$

where  $n = N_x + N_z + N_y - 3$ , and  $i = 1, 2, \dots, r$ .

Similarly the unbiased estimate of the covariance,  $i \neq j$ , is

$$a_{ij} = a_{ji} = \frac{1}{n} \left[ \sum_{k=1}^{N_x} (X_{ik} - \bar{X}_i)(X_{jk} - \bar{X}_j) + \sum_{k=1}^{N_z} (Z_{ik} - \bar{Z}_i)(Z_{jk} - \bar{Z}_j) + \sum_{k=1}^{N_y} (Y_{ik} - \bar{Y}_i)(Y_{jk} - \bar{Y}_j) \right],$$

where  $n = N_x + N_z + N_y - 3$ , and  $i, j = 1, 2, 3 \dots r, i \neq j$ .

In the present case  $N_y$  is 1, so that  $\sum_{k=1}^{N_x} (Y_{ik} - Y_i)^2$  and  $\sum_{k=1}^{N_z} (Y_{ik} - \bar{Y}_i) \cdot (Y_{jk} - \bar{Y}_j)$  each has the numerical value of zero. Consequently the determinant in this situation has the same value as the determinant found for

Problem I; similarly the co-factors will be numerically equal to those in the preceding problem.

It is now necessary to test the significance of the difference between the scores of an individual and each of the corresponding mean scores of the two populations with which the individual is being compared.

Let  $d_i = (\bar{Y}_i - \bar{X}_i) - (\bar{Z}_i - \bar{Y}_i)$ , and  $1/N = 4/N_v + 1/N_z + 1/N_s$ .

Define

$$V = \frac{1}{a} \left[ N \left( \sum_{i=1}^r \sum_{j=1}^r A_{ij} d_i d_j \right) \right]$$

and

$$F = \left[ \frac{N_z + N_v + N_s - 2 - r}{r} \right] (V)$$

with  $n_1 = (r)$  d.f. for the larger variance and  $n_2 = (N_z + N_v + N_s - 2 - r)$  d.f. for the smaller variance.

The  $F$  tables indicate the level of significance with which an individual can be classified into one of the two populations under consideration.

*Problem III. The Classification of an Individual into One of Two Populations.*

Having determined that an individual can be classified into either of two populations, it may then be desirable to determine into which population he is most appropriately placed. This involves the computation of the *generalized distance function*,  $D^2$ , from the values of the co-factors and determinant already available from the calculations in Problem II. Let the populations under consideration be  $\Pi_z$ ,  $\Pi_v$ , and  $\Pi_s$ , where  $\Pi_v$  is the individual to be classified. The generalized distance function is computed for  $\Pi_v$  and  $\Pi_z$ , and for  $\Pi_v$  and  $\Pi_s$ .

Let  $d_i = \bar{X}_i - \bar{Y}_i$

and 
$$D_{z,v}^2 = \frac{1}{a} \left[ \sum_{i=1}^r \sum_{j=1}^r A_{ij} d_i d_j \right].$$

Similarly, let  $d_i = \bar{Z}_i - \bar{Y}_i$

and 
$$D_{s,v}^2 = \frac{1}{a} \left[ \sum_{i=1}^r \sum_{j=1}^r A_{ij} d_i d_j \right].$$

The individual is then classified into that population for which  $D^2$  is smaller.

*Illustrative Problems*

*Problem I.*

The Heston Personal Adjustment Inventory was given to 145 female students training for the elementary teacher's credential at the University of

California and 100 women training for the general secondary credential. Three of the subtests of the Heston Inventory were selected for study, viz., Emotional Stability, Confidence, and Personal Relations. The hypothesis to be tested is that the two groups arose from the same population.

The statistics required for the calculation of  $T^2$  are shown in Table 1, which includes the sums, the sums of squares, and the sums of cross-products for each group.

TABLE 1  
Sums, Sums of Squares, and Sums of Cross-products

$\Sigma X_{1k} = 3267$	$\Sigma Z_{1k} = 4397$	$\Sigma X_{1k}X_{2k} = 104027$
$\Sigma X_{2k} = 3076$	$\Sigma Z_{2k} = 4194$	$\Sigma X_{1k}X_{3k} = 99335$
$\Sigma X_{3k} = 2979$	$\Sigma Z_{3k} = 4182$	$\Sigma X_{2k}X_{3k} = 93606$
$\Sigma X_{1k}^2 = 111609$	$\Sigma Z_{1k}^2 = 143269$	$\Sigma Z_{1k}Z_{2k} = 134181$
$\Sigma X_{2k}^2 = 99332$	$\Sigma Z_{2k}^2 = 129748$	$\Sigma Z_{1k}Z_{3k} = 131567$
$\Sigma X_{3k}^2 = 90803$	$\Sigma Z_{3k}^2 = 124768$	$\Sigma Z_{2k}Z_{3k} = 125150$
$N_x = 100$	$N_z = 145$	

The computation of  $T^2$  from the above data is next illustrated.

- a. Determine the values of  $d_i$ :

$$d_1 = 18.046932; d_2 = 14.123455; d_3 = 7.297829.$$

- b. Compute the numerical values of the estimates of variances and co-variances:

$$a_{11} = 60.945990$$

$$a_{22} = 54.133356$$

$$a_{12} = 43.356566$$

$$a_{23} = 25.354795$$

$$a_{13} = 27.829322$$

$$a_{33} = 25.563546$$

- c. The expression for the determinant becomes:

$$a = \begin{vmatrix} 60.945990 & 43.356566 & 27.829322 \\ 43.356566 & 54.133356 & 25.354795 \\ 27.829322 & 25.354795 & 25.563546 \end{vmatrix}$$

- d. Calculate the values of the co-factors:

$$A_{11} = 740.974907$$

$$A_{22} = 783.524456$$

$$A_{12} = 402.740815$$

$$A_{23} = 338.689246$$

$$A_{13} = -407.197752$$

$$A_{33} = 1419.419158$$

- e. Compute the value of the determinant:

$$a = 16365.953188$$

- f. The value of  $T^2$  may now be determined:

$$T^2 = 39.171612$$

- g. Calculate the probability that a  $T^2$  of this magnitude or larger could have arisen by chance if the two groups had arisen from the same population:

$$x = \frac{1}{1 + T^2/n} = .861178$$

Enter Pearson's Incomplete Beta Function tables with  $p = 120.5$ ,  $q = 1.5$ . If Pearson's tables are not available, it is possible to determine the significance of  $T^2$  by means of the  $F$  tables:

$$F = 12.9497, \text{ with } n_1 = 3, \text{ and } n_2 = 241.$$

An  $F$  of 3.88, with  $n_1 = 3$ ,  $n_2 = 200$ , is significant at the .01 level.

*Problem II.*

Can the individual  $Y_1 = 35$ ,  $Y_2 = 29$ ,  $Y_3 = 28$  be classified into either of the two populations of Problem I?

- a. Calculation of the values of  $d_i$ :

$$d_1 = 7.005862$$

$$d_2 = -1.684138$$

$$d_3 = -2.631379$$

- b. Compute the value of  $N$ :

$$N = .248948$$

- c. Determine the value of  $V$ :

$$V = .774669$$

- d. Calculate the value of the variance ratio ( $F$ ):

$$F = 62.2317$$

- e. Determine the significance of  $F = 62.2317$ , with  $n_1 = 3$ , and  $n_2 = 241$ . This value of  $F$  is significant at better than the .01 level. We conclude that the individual can be classified into one of these two populations.

*Problem III.*

Into which population of Problem I is the individual  $Y_1 = 35$ ,  $Y_2 = 29$ ,  $Y_3 = 28$  best classified?

- a. Determine the values of  $d_i$  for  $\Pi_x$ ,  $\Pi_y$ :

$$d_1 = -2.33$$

$$d_2 = 1.76$$

$$d_3 = 1.79$$

- b. Compute  $D_{x..xy}^2$ :

$$D_{x..xy}^2 = .80809$$

- c. Calculate the values of  $d_i$  for  $\Pi_x$ ,  $\Pi_y$ :

$$d_1 = -4.675862$$

$$d_2 = -.075862$$

$$d_3 = .841379$$

- d. Compute  $D_{x..xy}^2$ :

$$D_{x..xy}^2 = 1.262147$$

- e. Since  $D_{x..xy}^2$  is less than  $D_{y..xy}^2$ , we conclude that the individual ( $Y_1 = 35$ ,  $Y_2 = 29$ ,  $Y_3 = 28$ ) is best classified with  $\Pi_x$ .

*Concluding Statement*

The purpose of the present paper was to present procedures for answering three types of questions:

(1) Could two groups of subjects have arisen from the same population? In case the answer to the preceding question is negative, a second question arises:

(2) Is it possible to classify a given individual into either of the two populations mentioned in the first question?

If an affirmative answer is obtained, then a third question arises:

(3) Into which of the two populations is the individual best classified?

The formulas necessary to the solution of each question have been presented. An application of these methods of multivariate analysis has been made to a set of data about students preparing for the elementary and the secondary teaching credential.

#### REFERENCES

1. Fisher, R. A. The use of multiple measurements in taxonomic problems. *Ann. Eug.*, 1936-37, 7, 179-188.
2. Fisher, R. A. The statistical utilization of multiple measurements. *Ann. Eug.*, 1937-38, 8, 376-386.
3. Fisher, R. A. Contributions to mathematical statistics. New York: John Wiley and Sons, 1950.
4. Garrett, Henry E. The discriminant function and its use in psychology. *Psychometrika*, 1943, 8, 65-79.
5. Hodges, Joseph L., Jr. School of aviation medicine. Discriminatory analysis. I. Survey of discriminatory analysis. Project No. 21-49-004, Report No. 1. U.S.A.F. School of Aviation Medicine, Randolph Field, Texas, October, 1950.
6. Hotelling, H. The generalization of "Student's" ratio. *Ann. math. Stat.*, 1931, 2, 360-375.
7. Kendall, Maurice G. The advanced theory of statistics (2 Vols.). London: Charles Griffin and Co., Ltd., 1946.
8. Lindquist, E. F. Statistical analysis in educational research. New York: Houghton Mifflin Co., 1940.
9. Mahalanobis, P. C. On the generalized distance function in statistics. *Proc. nat. Inst. Sci. Ind.*, 1936, 12, 49-55.
10. Pearson, Karl. Incomplete beta function. Cambridge: Univ. Cambridge Press, 1948 (Rev. Ed.).
11. Rao, C. Radhakrishna. Tests with discriminant functions in multivariate analysis. *Sankhya: The Indian Journal of Statistics*, 1946, 7, 407-414.
12. Rao, C. Radhakrishna. The utilization of multiple measurements in problems of biological classification. *J. roy. Stat. Soc., Supp.*, 1948, 10, 159-203.
13. Rao, C. Radhakrishna. Tests of significance in multivariate analysis. *Biom.*, 1948, 35, 58-79.
14. Rao, C. Radhakrishna, and Slater, Patrick. Multivariate analysis applied to differences between neurotic groups. *Brit. J. Psychol.*, 1949, 2, (Part I), 17-29.

*Manuscript received 9/12/51*

*Revised manuscript received 11/30/51*

# NOTES ON A PROBLEM OF MULTIPLE CLASSIFICATION\*

FREDERIC M. LORD

EDUCATIONAL TESTING SERVICE

A solution is developed in implicit form for the problem of assigning  $N$  men to  $n$  jobs, the proportion of men to be assigned to each job being specified in advance.

Suppose that it is desired to assign  $N$  men to  $n$  jobs, the proportion of men to be assigned to each job being specified in advance. It is desired to maximize the average weighted productivity of the men, the productivity of each man being weighted according to the importance of the job to which he is assigned. It is assumed that the productivity of each man for each job is known in advance and can be used as a basis for assignment. If  $x_{ia}$  is the productivity of man  $a$  for job  $i$ , we can indicate the productivity of all men assigned to job  $i$  by  $\sum^* x_{ia}$ . The quantity to be maximized is then

$$Q = w_1 \sum^* x_{1a} + w_2 \sum^* x_{2a} + \cdots + w_n \sum^* x_{na}, \quad (1)$$

where  $w_i$  is the weight assigned to job  $i$ .

A solution to an almost identical problem, assuming all weights to be unity, has been given by Hubert E. Brogden (1); these and related problems have more recently been treated by Thorndike (2); Votaw (3) has quite recently made important contributions toward a rapidly converging successive approximation method for the practical solution of such problems.† The present paper is primarily concerned with developing in analytic form an implicit solution that is effectively the same as Brogden's; it is not immediately concerned with the problem of obtaining a practical solution by successive approximations.

## *The Two-Dimensional Case*

Consider first the case when  $n = 2$ . The contribution of any individual to  $Q$  will be  $w_1 x_{1a}$  if he is assigned to job 1,  $w_2 x_{2a}$  if he is assigned to job 2. The differential effect of the job assignment on  $Q$  is in this case  $w_1 x_{1a} - w_2 x_{2a}$ . All individuals for whom the quantity  $w_1 x_{1a} - w_2 x_{2a}$  has a given value are interchangeable with each other in their effect upon  $Q$ , and consequently are

\*The author wishes to thank Dr. Hubert Brogden and Dr. Paul Horst for their helpful discussion and criticism.

†Dr. Paul S. Dwyer has recently developed important practical methods of solution, as yet unpublished.

interchangeable with each other for purposes of job assignment. Furthermore, if certain individuals characterized by some specified value of  $w_1x_{1a} - w_2x_{2a}$  are properly assigned to job 1, then all individuals with higher values must also be assigned to job 1. The same relation obtains between  $w_2x_{2a} - w_1x_{1a}$  and assignment to job 2.

Suppose a scatter diagram is plotted with  $x_1$  and  $x_2$  as axes, each individual being represented by a point corresponding to his values of  $x_1$  and  $x_2$ . The reasoning just given shows that the optimum assignment of people to jobs corresponds to the division of the scatter diagram into two regions by the line  $w_1x_1 = w_2x_2 + k$ , where  $k$  is a constant to be determined so that the required proportions of individuals fall in the two regions created.

We thus have the result for  $n = 2$  that the optimum assignment corresponds to a region bounded by a straight line with a slope of  $w_2/w_1$ . This conclusion holds irrespective of the shape of the bivariate frequency distribution represented by the scatter diagram. The intercept of the line must be determined so that the proportions of cases cut off by the line are equal to the predetermined proportions of people to be assigned to the two jobs.

#### *The Three-Dimensional Case*

Let us next consider the case where  $n = 3$ . Using  $x_1$ ,  $x_2$ , and  $x_3$  as axes, a three-dimensional scatter plot may be prepared representing the data. The region of space containing the individuals to be assigned to job 1 will necessarily include all positive values on the  $x_1$  axis above a certain point. Consider the surface bounding this region (region 1) from the region containing the individuals to be assigned to job 2 (region 2).

Suppose that the desired regions have already been set up. There will be of necessity certain individuals who could equally well be assigned to jobs 1 or 2, but who should not be assigned to job 3. Such individuals lie on the boundary surface between jobs 1 and 2. For such people, the  $x_3$  score can have no effect on the decision as to whether they should be assigned to job 1 or to job 2. Consequently, the boundary between region 1 and region 2 is parallel to the  $x_3$  axis. By the same reasoning used for the case where  $n = 2$ , it follows that this boundary must be represented by the equation  $w_1x_1 = w_2x_2 + k_{12}$  where  $k_{12}$  is a constant to be determined so as to obtain the proper proportion of people in each region. Similarly the boundaries between regions 1 and 3 and between regions 2 and 3 are respectively  $w_1x_1 = w_3x_3 + k_{13}$  and  $w_2x_2 = w_3x_3 + k_{23}$ . In the present case these three equations define 3 planes, each of which is functionally independent of one of the variables and consequently parallel to the corresponding coordinate axis.

Let us see if any further conditions should be imposed on these planes. Let  $p_i$  be the proportion of cases to be assigned to job  $i$ . Since  $\sum^n p_i = 1$ , only  $n - 1 = 2$  of the values of  $p_i$  can be determined arbitrarily. Since the values of the  $k$ 's must be adjusted so that each region contains the proper

proportion of cases, not more than one restriction can be imposed on the three  $k$ 's. It will now be shown that one restriction must be imposed.

Consider the augmented matrix representing the three planes:

$$\left\| \begin{array}{cccc} w_1 & -w_2 & 0 & k_{12} \\ w_1 & 0 & -w_3 & k_{13} \\ 0 & w_2 & -w_3 & k_{23} \end{array} \right\|.$$

If this matrix has a rank of 3, the three planes will intersect in only one point, and hence will define  $2^3 = 8$  regions. Since we require only three regions, we would have to select some combination of these eight regions that would form the three regions required. Furthermore, symmetry requires that each of the three regions be defined with respect to the three coordinate axes in some way that is symmetrical, as far as the axes are concerned, with the definitions for the other two regions—e.g., if one boundary is a half-plane, all boundaries must be half-planes; if two half-planes intersect each other in a line, all half-planes must intersect all others in a line, etc. Since it is manifestly impossible to form three regions from the eight available so as to meet this condition, it follows that the rank of the matrix must be less than 3.

This same conclusion may be reached via an alternative argument. There must be, at least theoretically, certain individuals who could be assigned to any one of the three jobs without changing the value of  $Q$ —such individuals will lie at the mutual intersection of the three boundary planes. Moreover, such individuals may vary in at least one dimension of our coordinate space, since their contribution to  $Q$  (irrespective of the job to which they are assigned) may vary in one dimension, i.e., may be large, small, or intermediate in value. These cases therefore cannot all be located at a single point in our coordinate space, but must be distributed at least over a line. Since all these cases lie at a place where all three regions are in simultaneous contact, it follows that the planes bounding the regions must all have at least one line in common. This requires that the rank of the matrix given must be less than 3.

Since the  $w$ 's are predetermined, and since one and only one condition can be imposed on the  $k$ 's, it follows that the rank of the matrix is exactly 2, and hence that the planes have only a straight line in common.

In summary, for the case when  $n = 3$ , the boundaries of the three regions are the three planes represented by the matrix given, with the additional condition that the  $k$ 's must be so chosen as to give the matrix a rank of 2. Each plane is parallel to one coordinate axis and intersects the other two planes in a straight line. The slopes of the planes are determined by the pre-assigned weights and are independent of the distribution of cases in the scatter plot. The intercepts of the planes must be determined so that the planes cut off the desired proportion of the cases.

Since three planes intersecting in the same straight line define six regions,

we must form our three regions from these six. If we denote the plane that separates region  $i$  from region  $j$  by  $b_{ij}$ , we find, for example, that there are at least two separate regions lying between  $b_{12}$  and  $b_{13}$ . We must choose as region 1 the one of these that includes the end of the  $x_1$  coordinate axis representing large positive values of  $x_1$ . This is equivalent to saying that the boundaries of the three regions are three half-planes, each terminating in the same straight line that is its common intersection with the other two half-planes.

#### *The n-Dimensional Case*

These conclusions may readily be extended to any number of dimensions. In  $n$  dimensions, each of the  $n$  regions will be in contact with each of the other regions. The boundary between two such regions will be the  $(n - 1)$ -space represented by the equation  $w_i x_i = w_j x_j + k_{ij}$ . This boundary is parallel to all coordinate axes except  $x_i$  and  $x_j$ . There will be  $\frac{1}{2}n(n - 1)$  such boundaries, corresponding to the  $\frac{1}{2}n(n - 1)$  possible pairs of regions. All the boundaries will have as their mutual intersection a single straight line. The rank of the augmented matrix representing the boundaries will be  $n - 1$ .

Since all  $(n - 1)$ -spaces acting as boundaries pass through the same straight line, it is possible to make a translation such that they will all pass through the origin. The equations for the boundaries can thus be written  $w_i(x_i - b_i) = w_j(x_j - b_j)$  or  $w_i x_i = w_j x_j + a_i - a_j$ , where the  $a$ 's (or  $b$ 's) are constants replacing the  $k$ 's and determining the intercepts of the  $(n - 1)$ -spaces. In all the preceding formulations  $k_{ij}$  may simply be replaced by  $a_i - a_j$ . With this formulation it is seen that we need determine only the  $n$  unknown constants  $a_1, a_2, \dots, a_n$  instead of the  $\binom{n}{2}$  values of  $k_{11}, k_{12}, \dots, k_{22}, k_{23}, \dots, k_{n-1,n}$ .

For  $n = 4$ , for example, the boundaries are represented by

$$\begin{vmatrix} w_1 & -w_2 & 0 & 0 & a_1 - a_2 \\ w_1 & 0 & -w_3 & 0 & a_1 - a_3 \\ w_1 & 0 & 0 & -w_4 & a_1 - a_4 \\ 0 & w_2 & -w_3 & 0 & a_2 - a_3 \\ 0 & w_2 & 0 & -w_4 & a_2 - a_4 \\ 0 & 0 & w_3 & -w_4 & a_3 - a_4 \end{vmatrix},$$

where the  $a$ 's are restricted by the fact that the rank of the matrix must be 3.

#### *Formulation in Terms of Definite Integrals*

If the form of the multivariate scatter plot can be specified analytically by the continuous frequency function  $f(x_1, x_2, \dots, x_n)$ , the proportion of cases in region  $i$  can be expressed as a multiple integral, as follows:

$$\begin{aligned}
 p_i &= \int_{-\infty}^{\infty} \int_{-\infty}^{(w_i x_i - a_i + a_n)/w_n} \int_{-\infty}^{(w_i x_i - a_i + a_{n-1})/w_{n-1}} \\
 &\quad \cdots \int_{-\infty}^{(w_i x_i - a_i + a_{i+1})/w_{i+1}} \int_{-\infty}^{(w_i x_i - a_i + a_{i-1})/w_{i-1}} \\
 &\quad \cdots \int_{-\infty}^{(w_i x_i - a_i + a_1)/w_1} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_{i-1} dx_{i+1} \cdots dx_n dx_i.
 \end{aligned} \tag{2}$$

If  $n = 4$ , the simultaneous equations to be solved for the  $a$ 's are

$$\begin{aligned}
 p_1 &= \int_{-\infty}^{\infty} \int_{-\infty}^{(w_1 x_1 - a_1 + a_4)/w_4} \int_{-\infty}^{(w_1 x_1 - a_1 + a_3)/w_3} \int_{-\infty}^{(w_1 x_1 - a_1 + a_2)/w_2} \\
 &\quad \cdot f(x_1, x_2, x_3, x_4) dx_2 dx_3 dx_4 dx_1, \\
 p_2 &= \int_{-\infty}^{\infty} \int_{-\infty}^{(w_2 x_2 - a_2 + a_1)/w_1} \int_{-\infty}^{(w_2 x_2 - a_2 + a_4)/w_4} \int_{-\infty}^{(w_2 x_2 - a_2 + a_3)/w_3} \\
 &\quad \cdot f(x_1, x_2, x_3, x_4) dx_3 dx_4 dx_1 dx_2, \\
 p_3 &= \int_{-\infty}^{\infty} \int_{-\infty}^{(w_3 x_3 - a_3 + a_2)/w_2} \int_{-\infty}^{(w_3 x_3 - a_3 + a_1)/w_1} \int_{-\infty}^{(w_3 x_3 - a_3 + a_4)/w_4} \\
 &\quad \cdot f(x_1, x_2, x_3, x_4) dx_4 dx_1 dx_2 dx_3, \\
 p_4 &= \int_{-\infty}^{\infty} \int_{-\infty}^{(w_4 x_4 - a_4 + a_3)/w_3} \int_{-\infty}^{(w_4 x_4 - a_4 + a_2)/w_2} \int_{-\infty}^{(w_4 x_4 - a_4 + a_1)/w_1} \\
 &\quad \cdot f(x_1, x_2, x_3, x_4) dx_1 dx_2 dx_3 dx_4.
 \end{aligned} \tag{3}$$

The reasoning by which these integrals are written down may be illustrated by the first multiple integral given for the case when  $n = 4$ . We wish to find the frequency in region 1, which is known to be bounded by the three hypersurfaces  $w_2 x_2 = w_1 x_1 - a_1 + a_2$ ,  $w_3 x_3 = w_1 x_1 - a_1 + a_3$ , and  $w_4 x_4 = w_1 x_1 - a_1 + a_4$ . Suppose we wish to integrate first with respect to  $x_2$ . The last two of the hypersurfaces listed are parallel to the  $x_2$  axis, so they cannot serve as limits of integration and may be left out of consideration. Since the hypersurfaces can provide only one limit of integration, the other limit must be either  $+\infty$  or  $-\infty$ , but all points for which  $x_2 = +\infty$  are included in region 2. Hence the limits of integration with respect to  $x_2$  are  $-\infty$  and  $(w_1 x_1 - a_1 + a_2)/w_2$ . Similar reasoning will give us the limits of integration with respect to  $x_3$  and  $x_4$ . When these three integrations have been accomplished, there remains only one variable in our integrand—all the frequencies have been summed and expressed as a function of a single variable,  $x_1$ . Since, in the general case, there will be some finite frequency corresponding to every possible value of  $x_1$ , we must integrate from  $-\infty$  to  $+\infty$  with respect to this variable. The other multiple integrals given may be written down by similar lines of reasoning.

Since when  $n = 4$  the  $a$ 's must be so determined that the matrix of the boundary surfaces has a rank of 3, and since  $\sum p_i = 1$ , any three of the integral equations given is sufficient to determine the  $a$ 's. In practice, the solution of these equations will be extremely difficult in many cases. If  $f(x_1, x_2, \dots, x_n)$  is taken as the normal multivariate function, the necessary expressions at present available for the multiple integrals are too cumbersome to be of practical use. Unless simpler expressions can be found, the iterative methods developed by Brogden and Votaw will probably be found to be the most satisfactory method for handling actual numerical problems.

*The Case where Some Individuals Are to be Rejected*

We may consider briefly the additional case where  $p_0$  of the individuals

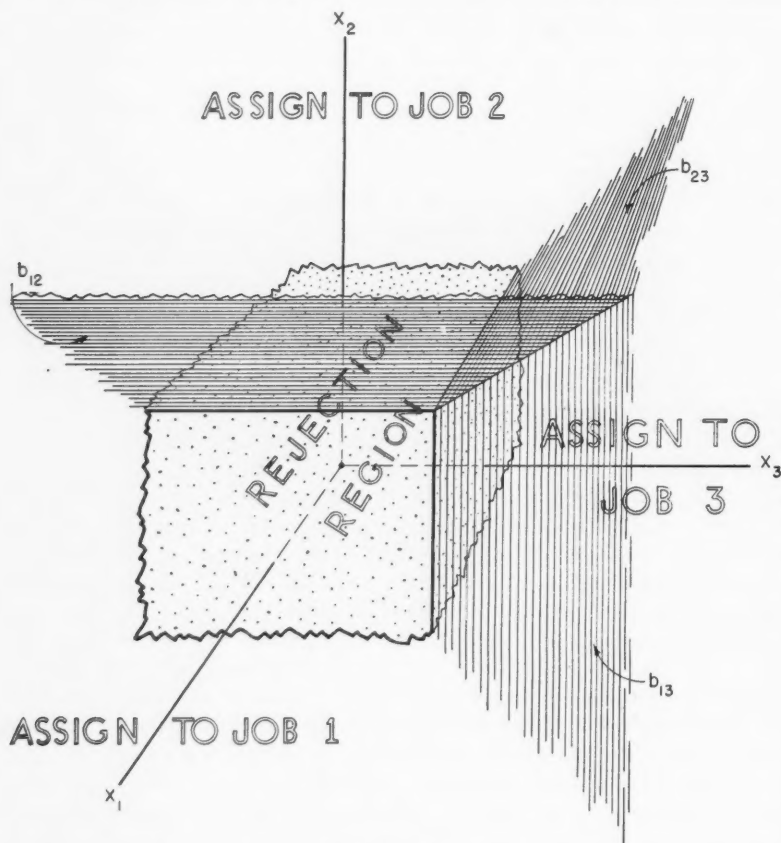


FIGURE 1  
Illustrative Assignment Regions and Rejection Region for a Three-job Problem

are not to be assigned to any job, but are to be rejected entirely. Now if individuals in region  $i$  are to be rejected, they must be rejected solely on the basis of  $x_i$ , irrespective of their score on other variables. Otherwise some people who were included in region  $i$  would have lower values of  $x_i$  than those of some people who were rejected. Consequently, some hyperplane of the form  $x_i = c_i$ , where  $c_i$  is an unknown constant, must be the boundary ( $b_{i0}$ ) between region  $i$  and the rejection region.

Next consider an individual who lies exactly on the boundary between region  $i$  and region  $j$ —this individual may equally well be assigned to either job  $i$  or job  $j$ . If this individual has a low value of  $x_i$ , he may simultaneously lie exactly on the boundary between region  $i$  and the rejection region, in which case he also may equally well be assigned to job  $i$  or rejected entirely. It follows that this individual, and all similar individuals, may equally well be rejected or assigned to job  $j$  and that consequently they all lie exactly on the boundary between region  $j$  and the rejection region. We have thus proved that the intersection of  $b_{i0}$  with  $b_{ij}$  coincides with the intersection of  $b_{i0}$  with  $b_{j0}$ . Since all boundaries between jobs intersect in a single straight line, the boundaries of the rejection region must intersect each other in a point lying on this line.

The rejection region is thus a rectangular region each of whose sides is parallel to  $n - 1$  of the coordinate axes. It obviously includes all large negative values of all the variables. (In practice negative values of  $x_i$  may not occur, but we need not exclude the possibility of their occurrence in our theoretical discussion.) The intersection of any two sides of this region coincides with their intersection with one of the boundaries between jobs. This may be visualized for the three-dimensional case as follows (see Figure 1): The rejection region is a rectangular solid containing all large negative values of all variables. Only the three upper faces of this rectangular solid can be visualized, since the other three faces are at  $-\infty$ . Each face is perpendicular to one of the coordinate axes. An oblique line having a positive slope with respect to all axes extends in a positive direction from the corner of the rectangular solid. The boundaries separating the three jobs are three oblique planes each of which joins this line with one of the edges of the rectangular solid and extends to  $\infty$  in the remaining directions.

The analytic equations corresponding to this situation may be readily written down. For the general case the integrals will be the same as given before in equation 2, with the exception that the lower limit of the first integral sign will be  $c_i$  instead of  $-\infty$ . In addition to the  $n$  equations of the type given, there will be an  $(n + 1)$ -th equation, as follows:

$$p_0 = \int_{-\infty}^{c_1} \int_{-\infty}^{c_2} \cdots \int_{-\infty}^{c_n} f(x_1, x_2, \cdots, x_n) dx_n \cdots dx_2 dx_1. \quad (4)$$

In addition to the restrictions previously imposed on the  $a$ 's, we now must impose  $n - 1$  restrictions on the  $c$ 's corresponding to the fact that all

the rejection-region boundaries intersect in a single point through which all the other boundaries pass. These restrictions may be written down as follows:

$$w_j c_i - w_i c_j = a_i - a_j \quad (i, j = 1, \dots, n). \quad (5)$$

Only  $n - 1$  of the equations in (5) are independent.

#### REFERENCES

1. Brogden, H. E. An approach to the problem of differential prediction. *Psychometrika*, 1946, 11, 139-154.
2. Thorndike, R. L. The problem of classification of personnel. *Psychometrika*, 1950, 15, 215-236.
3. Votaw, D. F., Jr. Methods of solving some personnel-classification problems. *Psychometrika*, 1952, 17, 255-266.

*Manuscript received 10/22/51*

*Revised manuscript received 1/14/52*

## THE ROLLING TOTALS METHOD OF COMPUTING SUMS, SUMS OF SQUARES, AND SUMS OF CROSS-PRODUCTS

J. DOUGLAS AYERS,

DEFENCE RESEARCH MEDICAL LABORATORIES

AND

J. PERHAM STANLEY,

ABITIBI POWER AND PAPER COMPANY

The rolling totals method of computing sums, sums of squares, and sums of cross-products appears to have several advantages over usual methods, in that it saves time and requires less equipment. Only an IBM sorter and a tabulator equipped with Card Cycle Total Transfer Device are used. The method provides an immediate independent visual check on the accuracy of the sums of cross-products of each successive variable as it is run. Since controlling is done by sorting, the necessity for re-wiring after each run on the tabulator is eliminated. The wiring, illustrated by a diagram, is simple and straightforward.

### *Introduction*

A great deal of time is consumed in obtaining the sums, sums of squares, and sums of cross-products before correlational analysis between measures can be computed. A review of the literature indicates that there are several methods by which the total of products or squares is obtained without carrying out the individual calculations of products and squares. Such methods as digitizing, progressive digitizing, and digitizing without sorting have been discussed (1, 2, 3, 4). In general, one run through of the tabulator is required for each position of the multiplier. The partial totals thus obtained must then be added to secure the final totals. More recently, with the introduction of the electronic computing punch, it has become possible to obtain the summed products of variables up to a total of twenty digits on a trailer card (5). However, this method requires considerable time and is more subject to error than the methods mentioned above.

The rolling totals method described below appears to have several distinct advantages over those methods previously described. The main advantages are:

1. It saves considerable time and eliminates the need for a summary punch. Although the rolling totals method gives up counter space used in other methods for obtaining progressive totals for more variables, it still saves time by omitting three usual procedures. In this method there is no change of

wiring within each set of variables. The number of card cycles is greatly reduced because there are no intermediate totals printed and no summary cards punched. Finally, the process of summing the summary cards, with all that it entails, is eliminated.

2. This method provides immediate independent visual checks on the accuracy of the sums of cross-products of each successive variable as it is run. Generally speaking, the fewer steps involved in a computational process, the less chance of error. After the second run there is an immediate visual check on the cross-products  $X_1X_2$  and  $X_2X_1$ . The only possible errors, either in wiring the variables from the lower brushes or in a mis-sort, are caught immediately. The method has proved to be so accurate that a straight card count or an accumulating card count to obtain a check on the sums for each variable is actually unnecessary and takes up counter space that could be used for another variable.

3. Controlling is done by sorting so that wires need not be changed with each run through on the tabulator.

The values for  $\sum X_i^2$  and  $\sum X_iX_j$  are obtained directly and the final totals are printed on one sheet as a symmetrical matrix, with the cells of the principal diagonal containing the sums of squares and the remaining cells the sums of cross-products. Table 1 shows the algebraic set-up for  $n$  variables.

The values for  $\sum X_i$  are obtained by summing the detail cards on the tabulator without change of wiring. The basic principle for computing  $\sum X_i^2$  and  $\sum X_iX_j$  is that of (a) sorting digit cards after detail cards, and (b) summing the detail cards on the tabulator by the "rolling totals" method. This procedure is repeated for each variable.

### Equipment

To perform the operations described below, the only pieces of IBM equipment required are a sorter and a tabulator equipped with the Card

TABLE 1  
Algebraic Illustration of Sums Obtained by the Rolling Totals Method

	Variables					Sorted on
	$X_1$	$X_2$	$X_3$	— —	$X_n$	
Sums	[ $\sum X_1$	$\sum X_2$	$\sum X_3$	— —	$\sum X_n$ ]	none
Sums of	[ $\sum X_1^2$	$\sum X_1X_2$	$\sum X_1X_3$	— —	$\sum X_1X_n$ ]	$X_1$
Squares	$\sum X_2X_1$	$\sum X_2^2$	$\sum X_2X_3$	— —	$\sum X_2X_n$	$X_2$
and	$\sum X_3X_1$	$\sum X_3X_2$	$\sum X_3^2$	— —	$\sum X_3X_n$	$X_3$
Sums of	—	—	—	— —	—	—
Cross-	—	—	—	— —	—	—
Products	$\sum X_nX_1$	$\sum X_nX_2$	$\sum X_nX_3$	— —	$\sum X_n^2$ ]	$X_n$

Cycle Total Transfer Device, or "rolling" device for short.\* Tabulators may be equipped with this device, replacing progressive digiting switches which will then be unnecessary. Moreover, since no summary punch or other equipment is required, this method is more economical than others.

### *Procedure*

The sums of squares and cross-products are obtained by a single run of the cards through the tabulator for each variable. The cards are sorted together with a set of X-punched digit cards in descending order on the field of the multiplier, that is, the field to be squared. For a one-column multiplier, a maximum of nine digit cards (9 to 1), for two columns, ninety-nine cards (99 to 01), and for three columns, 999 cards (999 to 001) will be required. When the method is used a great deal, it is more efficient to have three sets of digit cards. The one-column set has 9's punched continuously in columns 1 to 79, another has 8's, another 7's, and so on through 1's. One of the two-column sets has 99's punched continuously, another 98's, another 97's, and so on through 01's in columns 1 to 78. The other two-column set starts at column 2 and goes through 79. All cards are X-punched in column 80.

Now if there are no detail cards for a number, at least a digit card for that number will be present. All digit cards larger than the largest number in any variable are removed as these are unnecessary. The file of cards is now sorted with the digit cards following the detail cards. The last card of the file will be a digit card numbered 1 or 01. The cards are tabulated on an accounting machine which is equipped with the rolling device on counters 2A, 2B, 4A, 4B, 6A, 6B, 8A, and 8B—the A-B counter group. The totals accumulated are rolled from the A-B counters into the C-D counters by the digit cards. When the last digit card passes through the tabulator, the total sums are printed.

### *Method*

Figure 1 shows the plugging for obtaining the sums, sums of squares, and sums of cross-products for six variables  $X_1$  to  $X_6$ . The top counters A-B are paired with the bottom counters C-D. By making all the A's and B's common and all the C's and D's common, a maximum of forty counting units is available in each group. The sums are rolled from the A-B counter group into the C-D counter group on each digit card. All counters are usually equipped with the Card Cycle Total Transfer Device, although only the A-B counters need be for the method proposed here.

The variables are plugged to the counter group A-B through the No-X side of Selectors E and F† in order to prevent accumulating from the digit

\*This device can be attached to type 405 and type 416 Accounting Machines. It comes as standard equipment on the new type 407. We used a type 416 tabulator and a type 080 sorter.

†Selectors E and F are not standard equipment on all machines. However, Selectors A, B, C, and D can be used. Selectors E and F were used because of convenience in wiring.

INTERNATIONAL BUSINESS MACHINES CORPORATION  
ALPHABETIC ACCOUNTING MACHINE CONTROL PANEL  
TYPE 405 WITH NET BALANCE COUNTERS

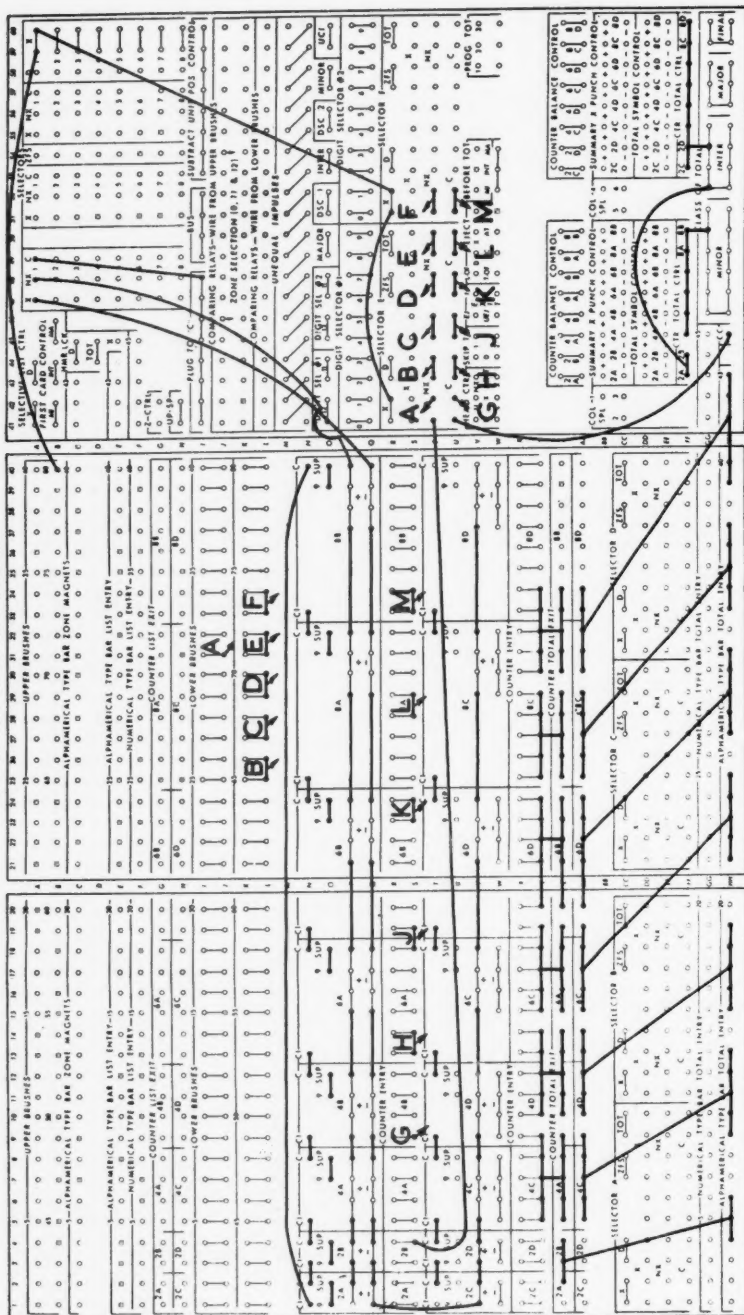


Figure 1  
Wiring Diagram for the Rolling Totals Method

cards. The add and subtract impulses are also under control of a selector, in this case Selector 1. When the digit card passes the upper brushes, the X in column 80 sets up Selectors E and F so that when this card passes the lower brushes, these selectors are in the controlled position and accumulation from the digit card is eliminated. When the digit card passes the lower brushes, counter group A-B subtracts and the totals standing in it are transferred to counter group C-D, which accumulates the totals from each digit card. After the transfer, counter group A-B will contain the same figures as before the transfer. If there are no detail cards but only a digit card present for a group, the totals transferred from the previous group will be transferred again for this group. After all the cards have passed through the machine, counter group C-D contains the sum of squares and sums of cross-products for the variable sorted. These totals are wired to print as final totals.

Card count is optional. It is wired in this illustration through Selector E as for the variables. Progressive accumulation was not wanted; therefore, counters 2A and 2B are wired directly to Type Bar Total Entry. Counter Total Control for card count is wired to Intermediate for printing on the same cycle as the final totals. If card count is wired the same as a variable and allowed to roll with the other totals, the result will be  $\sum X_i$  for the variable sorted. This provides an extra check on the accuracy of the work.

Table 2 shows an actual run by the rolling totals method with appropriate headings added. The column headings indicate the columns in which the

TABLE 2  
Sample of Sums, Sums of Squares and Sums of Cross-Products  
Obtained by the Rolling Totals Method

Variables	Card Count	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
Column	1	3	66	66	67	77	77
Headings	1	1	56	78	90	12	34
Sums	206	1027	2771	2628	2772	2760	2808
Sums of	206	5903	14484	13638	14448	14367	14517
Squares	206	14484	40775	38084	40106	39364	40158
and	206	13638	38084	37092	38316	37762	38597
Sums of	206	14448	40106	38316	41304	40088	40702
Cross-	206	14367	39364	37762	40088	40704	40332
Products	206	14517	40158	38597	40702	40332	42382

variables have been punched, and they are printed by running through the usual heading cards one at a time, followed by an X-punched card. The sums of the six variables are obtained by running the unsorted detail cards, followed

by a single digit card, through the tabulator. The matrix of the sums of squares and sums of cross-products is obtained by the method explained above—with only a single run of the sorted cards for each variable. At no time is it necessary to change the wiring because the summing is controlled by the sorting. Since row indication cannot be printed automatically, care must be taken to sort each variable in order.

#### *Final Comments*

Since the maximum number of counters available is usually forty,\* it will be possible to operate with ten variables if the sums of squares never exceed four digits, with eight variables for five digits, and so on. When all forty counters are required for the variables, card count can be omitted. If more than forty counters are needed, it will be necessary to do the runs in two or more parts. In this case, the partial matrices will be on two or more sheets which can be fitted together to make a complete symmetrical matrix.

An operator dealing with a minimum of about 200 cards in each of two or more decks can keep a tabulator running continuously. Therefore, if one is operating with only one set of cards, it is often more economical to gang punch a second deck and sort on alternate variables.

#### REFERENCES

1. Castore, George F., and Dye, William S., III. A simplified punch card method of determining sums of squares and sums of products. *Psychometrika*, 1949, **14**, 243-250.
2. Deemer, Walter L. (Editor). Records, analysis, and test procedures. AAF Aviation Psychology Program Research Reports, No. 18. Washington: U. S. Government Printing Office, 1947, 405-442.
3. Dwyer, Paul S., and Meacham, Allan D. The preparation of correlation tables on a tabulator equipped with digit selection. *J. Amer. statist. Ass.*, 1937, **32**, 654-662.
4. Ellis, Max E., and Riopelle, Arthur J. An efficient punched-card method of computing  $\Sigma X$ ,  $\Sigma X^2$ , and  $\Sigma XY$ , and higher moments. *Psychometrika*, 1948, **13**, 79-85.
5. International Business Machines Corporation. Calculating punch, type 602A, principles of operation. New York: IBM, 1948.

\*Actually fifty-six with the new type 407.

*Manuscript received 10/24/51*

*Revised manuscript received 12/31/51*

AN EXTENDED TABLE OF CHI-SQUARE FOR TWO DEGREES  
OF FREEDOM, FOR USE IN COMBINING PROBABILITIES  
FROM INDEPENDENT SAMPLES

MORDECAI H. GORDON, EDWARD H. LOVELAND, AND EDWARD E. CURETON  
VA HOSPITAL, CHILLICOTHE, OHIO; GEORGIA INSTITUTE OF TECHNOLOGY;  
AND UNIVERSITY OF TENNESSEE

A table of values of Chi-square for two degrees of freedom corresponding to values of  $P$  from .001 to .999 is presented, together with a description and an example of its use in combining probabilities from two or more independent samples to obtain an aggregate probability.

Duplicate studies, and studies bearing on the same essential problem but based on samples from non-identical populations, often lead to tests of significance whose associated probabilities are quite unequal. If each sample is small, moreover, the null hypothesis may not be rejected with any great certainty on the basis of any one study. But if the results of the several studies are all or nearly all in the same direction, there is additional evidence against the null hypothesis. In such cases we may wish to base a new test of significance on the combined data from all the studies.

Fisher (2) has described a method for combining the probabilities yielded by two or more studies, by making use of the fact that the distribution of the sum of several values of Chi-square is itself a Chi-square distribution, and that the value of Chi-square for two degrees of freedom is  $-2$  times the natural logarithm of the probability. To use this method it is necessary to find from each original study the exact probability corresponding to the statistic involved in the test of significance.

Since many of the commonly used tables do not give directly the natural logarithms of the numbers between 0.000 and 1.000, the preparation of an extended direct table of values of Chi-square appeared desirable. The accompanying table gives four-decimal values of Chi-square, for two degrees of freedom, corresponding to values of  $P$  from .001 to .999. These values were found by doubling the entries in a five-decimal table of natural logarithms (3), changing the signs, and rounding off to four decimals. For values of  $P$  greater than .05 and accurate to four or five decimals, linear interpolation will be accurate enough for most practical purposes. For values of  $P$  less than .05 and accurate to more than three decimals, the corresponding values of Chi-

TABLE 1  
Values of Chi-Square for Two Degrees of Freedom  
Corresponding to Values of P from .001 to .999.

P	0	1	2	3	4	5	6	7	8	9
.00		13.8155	12.4292	11.6183	11.0429	10.5966	10.2320	9.9237	9.6566	9.4211
.01	9.2103	9.0197	8.8457	8.6856	8.5374	8.3994	8.2703	8.1491	8.0348	7.9266
.02	7.8240	7.7265	7.6334	7.5445	7.4594	7.3778	7.2993	7.2238	7.1511	7.0809
.03	7.0131	6.9475	6.8840	6.8225	6.7628	6.7048	6.6485	6.5937	6.5403	6.4884
.04	6.4378	6.3884	6.3402	6.2931	6.2471	6.2022	6.1582	6.1152	6.0731	6.0319
.05	5.9915	5.9519	5.9130	5.8749	5.8375	5.8008	5.7648	5.7294	5.6946	5.6604
.06	5.6268	5.5938	5.5612	5.5292	5.4977	5.4667	5.4362	5.4061	5.3765	5.3473
.07	5.3185	5.2902	5.2622	5.2346	5.2074	5.1805	5.1540	5.1279	5.1021	5.0766
.08	5.0515	5.0266	5.0021	4.9778	4.9539	4.9302	4.9068	4.8837	4.8608	4.8382
.09	4.8159	4.7938	4.7719	4.7503	4.7289	4.7078	4.6868	4.6661	4.6456	4.6253
.10	4.6052	4.5853	4.5656	4.5461	4.5267	4.5076	4.4886	4.4699	4.4512	4.4328
.11	4.4145	4.3965	4.3785	4.3607	4.3431	4.3256	4.3083	4.2912	4.2741	4.2573
.12	4.2405	4.2239	4.2075	4.1911	4.1749	4.1589	4.1429	4.1271	4.1115	4.0959
.13	4.0804	4.0651	4.0499	4.0348	4.0198	4.0050	3.9902	3.9755	3.9610	3.9466
.14	3.9322	3.9180	3.9039	3.8898	3.8759	3.8620	3.8483	3.8346	3.8211	3.8076
.15	3.7942	3.7810	3.7677	3.7546	3.7416	3.7287	3.7158	3.7030	3.6903	3.6777
.16	3.6652	3.6527	3.6403	3.6280	3.6158	3.6036	3.5915	3.5795	3.5676	3.5557
.17	3.5439	3.5322	3.5205	3.5089	3.4974	3.4859	3.4745	3.4632	3.4519	3.4407
.18	3.4296	3.4185	3.4075	3.3965	3.3856	3.3748	3.3640	3.3533	3.3426	3.3320
.19	3.3215	3.3110	3.3005	3.2901	3.2798	3.2695	3.2593	3.2491	3.2390	3.2289
.20	3.2189	3.2089	3.1990	3.1891	3.1793	3.1695	3.1598	3.1501	3.1404	3.1308
.21	3.1213	3.1118	3.1023	3.0929	3.0836	3.0742	3.0650	3.0557	3.0465	3.0374
.22	3.0283	3.0192	3.0102	3.0012	2.9922	2.9833	2.9744	2.9656	2.9568	2.9481
.23	2.9394	2.9307	2.9220	2.9134	2.9049	2.8963	2.8878	2.8794	2.8710	2.8626
.24	2.8542	2.8459	2.8376	2.8294	2.8212	2.8130	2.8048	2.7967	2.7887	2.7806
.25	2.7726	2.7646	2.7567	2.7487	2.7408	2.7330	2.7252	2.7174	2.7096	2.7019
.26	2.6941	2.6865	2.6788	2.6712	2.6636	2.6561	2.6485	2.6410	2.6335	2.6261
.27	2.6187	2.6113	2.6039	2.5966	2.5893	2.5820	2.5747	2.5675	2.5603	2.5531
.28	2.5459	2.5388	2.5317	2.5246	2.5176	2.5105	2.5035	2.4965	2.4896	2.4827
.29	2.4757	2.4689	2.4620	2.4552	2.4484	2.4416	2.4348	2.4280	2.4213	2.4146
.30	2.4079	2.4013	2.3947	2.3880	2.3815	2.3749	2.3683	2.3618	2.3553	2.3488
.31	2.3424	2.3359	2.3295	2.3231	2.3167	2.3104	2.3040	2.2977	2.2914	2.2851
.32	2.2789	2.2726	2.2664	2.2602	2.2540	2.2479	2.2417	2.2356	2.2295	2.2234
.33	2.2173	2.2113	2.2052	2.1992	2.1932	2.1872	2.1813	2.1753	2.1694	2.1635
.34	2.1576	2.1517	2.1459	2.1400	2.1342	2.1284	2.1226	2.1169	2.1111	2.1054
.35	2.0996	2.0939	2.0882	2.0826	2.0769	2.0713	2.0656	2.0600	2.0544	2.0489
.36	2.0433	2.0378	2.0322	2.0267	2.0212	2.0157	2.0102	2.0048	1.9993	1.9939
.37	1.9885	1.9831	1.9777	1.9724	1.9670	1.9617	1.9563	1.9510	1.9457	1.9404
.38	1.9352	1.9299	1.9247	1.9194	1.9142	1.9090	1.9038	1.8987	1.8935	1.8884
.39	1.8832	1.8781	1.8730	1.8679	1.8628	1.8577	1.8527	1.8476	1.8426	1.8376

TABLE 1—Continued

P	0	1	2	3	4	5	6	7	8	9
.40	1.8326	1.8276	1.8226	1.8176	1.8127	1.8077	1.8028	1.7979	1.7930	1.7881
.41	1.7832	1.7783	1.7735	1.7686	1.7638	1.7590	1.7541	1.7493	1.7445	1.7398
.42	1.7350	1.7302	1.7255	1.7208	1.7160	1.7113	1.7066	1.7019	1.6973	1.6926
.43	1.6879	1.6833	1.6787	1.6740	1.6694	1.6648	1.6602	1.6556	1.6511	1.6465
.44	1.6420	1.6374	1.6329	1.6284	1.6239	1.6194	1.6149	1.6104	1.6059	1.6015
.45	1.5970	1.5926	1.5881	1.5837	1.5793	1.5749	1.5705	1.5661	1.5618	1.5574
.46	1.5531	1.5487	1.5444	1.5401	1.5357	1.5314	1.5271	1.5229	1.5186	1.5143
.47	1.5100	1.5058	1.5016	1.4973	1.4931	1.4889	1.4847	1.4805	1.4763	1.4721
.48	1.4679	1.4638	1.4596	1.4555	1.4513	1.4472	1.4431	1.4390	1.4349	1.4308
.49	1.4267	1.4226	1.4186	1.4145	1.4104	1.4064	1.4024	1.3983	1.3943	1.3903
.50	1.3863	1.3823	1.3783	1.3743	1.3704	1.3664	1.3624	1.3585	1.3545	1.3506
.51	1.3467	1.3428	1.3389	1.3350	1.3311	1.3272	1.3233	1.3194	1.3156	1.3117
.52	1.3079	1.3040	1.3002	1.2963	1.2925	1.2887	1.2849	1.2811	1.2773	1.2735
.53	1.2698	1.2660	1.2622	1.2585	1.2547	1.2510	1.2472	1.2435	1.2398	1.2361
.54	1.2324	1.2287	1.2250	1.2213	1.2176	1.2139	1.2103	1.2066	1.2030	1.1993
.55	1.1957	1.1920	1.1884	1.1848	1.1812	1.1776	1.1740	1.1704	1.1668	1.1632
.56	1.1596	1.1561	1.1525	1.1490	1.1454	1.1419	1.1383	1.1348	1.1313	1.1277
.57	1.1242	1.1207	1.1172	1.1137	1.1103	1.1068	1.1033	1.0998	1.0964	1.0929
.58	1.0895	1.0860	1.0826	1.0791	1.0757	1.0723	1.0689	1.0655	1.0621	1.0587
.59	1.0553	1.0519	1.0485	1.0451	1.0418	1.0384	1.0350	1.0317	1.0283	1.0250
.60	1.0217	1.0183	1.0150	1.0117	1.0084	1.0051	1.0018	.9985	.9952	.9919
.61	.9886	.9853	.9820	.9788	.9755	.9723	.9690	.9658	.9625	.9593
.62	.9561	.9528	.9496	.9464	.9432	.9400	.9368	.9336	.9304	.9272
.63	.9241	.9209	.9177	.9146	.9114	.9083	.9051	.9020	.8988	.8957
.64	.8926	.8895	.8863	.8832	.8801	.8770	.8739	.8708	.8677	.8646
.65	.8616	.8585	.8554	.8524	.8493	.8462	.8432	.8401	.8371	.8341
.66	.8310	.8280	.8250	.8220	.8189	.8159	.8129	.8099	.8069	.8039
.67	.8010	.7980	.7950	.7920	.7891	.7861	.7831	.7802	.7772	.7743
.68	.7713	.7684	.7655	.7625	.7596	.7567	.7538	.7508	.7479	.7450
.69	.7421	.7392	.7363	.7335	.7306	.7277	.7248	.7219	.7191	.7162
.70	.7133	.7105	.7076	.7048	.7020	.6991	.6963	.6934	.6906	.6878
.71	.6850	.6822	.6794	.6765	.6737	.6709	.6682	.6654	.6626	.6598
.72	.6570	.6542	.6515	.6487	.6459	.6432	.6404	.6377	.6349	.6322
.73	.6294	.6267	.6239	.6212	.6185	.6158	.6131	.6103	.6076	.6049
.74	.6022	.5995	.5968	.5941	.5914	.5887	.5861	.5834	.5807	.5780
.75	.5754	.5727	.5700	.5674	.5647	.5621	.5594	.5568	.5541	.5515
.76	.5489	.5462	.5436	.5410	.5384	.5358	.5331	.5305	.5279	.5253
.77	.5227	.5201	.5175	.5150	.5124	.5098	.5072	.5046	.5021	.4995
.78	.4969	.4944	.4918	.4892	.4867	.4841	.4816	.4791	.4765	.4740
.79	.4714	.4689	.4664	.4639	.4613	.4588	.4563	.4538	.4513	.4488

TABLE 1—Continued

P	0	1	2	3	4	5	6	7	8	9
.80	.4463	.4438	.4413	.4388	.4363	.4338	.4313	.4289	.4264	.4239
.81	.4214	.4190	.4165	.4140	.4116	.4091	.4067	.4042	.4018	.3993
.82	.3969	.3945	.3920	.3896	.3872	.3847	.3823	.3799	.3775	.3751
.83	.3727	.3703	.3678	.3654	.3630	.3606	.3583	.3559	.3535	.3511
.84	.3487	.3463	.3440	.3416	.3392	.3368	.3345	.3321	.3297	.3274
.85	.3250	.3227	.3203	.3180	.3156	.3133	.3110	.3086	.3063	.3040
.86	.3016	.2993	.2970	.2947	.2924	.2901	.2877	.2854	.2831	.2808
.87	.2785	.2762	.2739	.2716	.2693	.2671	.2648	.2625	.2602	.2579
.88	.2557	.2534	.2511	.2489	.2466	.2443	.2421	.2398	.2376	.2353
.89	.2331	.2308	.2286	.2263	.2241	.2219	.2196	.2174	.2152	.2129
.90	.2107	.2085	.2063	.2041	.2019	.1996	.1974	.1952	.1930	.1908
.91	.1886	.1864	.1842	.1820	.1798	.1777	.1755	.1733	.1711	.1689
.92	.1668	.1646	.1624	.1603	.1581	.1559	.1538	.1516	.1494	.1473
.93	.1451	.1430	.1408	.1387	.1366	.1344	.1323	.1301	.1280	.1259
.94	.1238	.1216	.1195	.1174	.1153	.1131	.1110	.1089	.1068	.1047
.95	.1026	.1005	.0984	.0963	.0942	.0921	.0900	.0879	.0858	.0837
.96	.0816	.0796	.0775	.0754	.0733	.0713	.0692	.0671	.0650	.0630
.97	.0609	.0589	.0568	.0547	.0527	.0506	.0486	.0465	.0445	.0424
.98	.0404	.0384	.0363	.0343	.0323	.0302	.0282	.0262	.0241	.0221
.99	.0201	.0181	.0161	.0140	.0120	.0100	.0080	.0060	.0040	.0020

square for two degrees of freedom may be found by multiplying their common logarithms by  $-4.60517$ .

To find the combined probability from  $n$  duplicate or parallel studies, find the value of Chi-square (for two degrees of freedom) corresponding to the *one-tail*  $P$  yielded by each study. If some of the studies yield results of opposite sign from the majority, the values of Chi-square to be used are those corresponding to  $(1 - P)$  rather than  $P$ . Add these Chi-square values. The sum may be treated as a Chi-square based on  $2n$  degrees of freedom, the  $P$  corresponding to this Chi-square being the probability for the combined significance test.

*Example:* McNemar and Terman (5) report data from studies by Thorndike based on the administration of a paper-and-pencil intelligence test (the CAVD) to similarly selected samples of boys and girls from 13 to 17 years of age. Each of thirteen age groups in various cities was treated as a separate sample. The critical ratios of the differences between pairs of standard deviations were computed by large-sample methods. The thirteen pairs of  $N$ 's, critical ratios, one-sided probabilities, and Chi-square values for two degrees of freedom are as follows:

Sample	N(boys)	N(girls)	CR	P	Chi-square
1	252	279	1.00	.159	3.6777
2	555	637	1.28	.100	4.6502
3	439	490	2.04	.021	7.7265
4	170	155	.29	.386	1.8987
5	220	274	.24	.405	1.8077
6	377	494	1.61	.054	5.8375
7	373	405	1.22	.111	4.3965
8	153	189	.33	.371	1.9831
9	180	198	.50	.309	2.3488
10	615	716	.66	.255	2.7330
11	775	1054	1.17	.121	4.2239
12	728	865	-.38	.648	.8677
13	385	436	.18	.429	1.6926
					43.7989

At 26 degrees of freedom, the value of the combined Chi-square is significant at the .02 level but not at the .01 level. This is a one-sided test, however, while the hypothesis (no difference in variability as between boys and girls) calls for a two-sided test. The tabled significance values must therefore be doubled, and we conclude that the hypothesis is to be rejected at the .04 level but not at the .02 level. The positive CR's in this example refer to pairs of samples in which the variability of boys exceeds that of girls. Earlier studies would have led the authors to state this as the alternative to the null hypothesis. In that case the one-sided test would apply, and the null hypothesis would be rejected at the .02 level but not at the .01 level.\*

It has been pointed out (4, 6) that when the P's to be combined come from four-fold tables or binomial distributions with some one or more frequencies small, this method of combining probabilities is not accurate. In such cases P can take only discrete values, and the neighboring values are not very close together. Fisher's method of combining probabilities applies strictly only when the original tests of significance are of such nature that all values of P are equally probable *a priori*. When values of P do come from discrete distributions, however, the use of Fisher's method is "conservative"; i.e., the null hypothesis will be rejected less often by this test than it will by a more exact test.

A convenient list of statistical tables, including most of the original tables which yield exact probability values from significance tests, has been given recently by Bancroft (1).

\*One reviewer of this paper suggests that in cases such as the one in this example, where the significance test is based on a *critical ratio* (which already assumes a normal sampling distribution), we may use the *t*-test of the hypothesis that the mean CR is zero, taking  $1/n$  times the variance of the CR's in the  $n$  samples as an estimate of the sampling variance of these means. The weighted mean CR in the present example is .822, the standard error of this mean is .202, and  $t = 4.07$ , which for 12 degrees of freedom is significant (in a two-sided test) at the .01 level but not at the .001 level. In this instance the *t*-test is more sensitive than is the Chi-square test based on the combined distribution.

## REFERENCES

1. Bancroft, T. A. Probability values for the common tests of hypotheses. *J. Amer. statist. Assoc.*, 1950, **45**, 211-217.
2. Fisher, R. A. Statistical methods for research workers, Tenth Ed. London: Oliver and Boyd, 1946. Pp. 99-101.
3. Hodgman, C. D., (compiler). Mathematical tables, Ninth Ed. Cleveland: Chemical Rubber Pub. Co., 1948. Pp. 168-169.
4. Lancaster, H. O. The combination of probabilities arising from data in discrete distributions. *Biometrika*, 1949, **36**, 370-382.
5. McNemar, Quinn, and Terman, L. M. Sex differences in variational tendency. *Genet. Psychol. Monog.*, 1936, **18**, No. 1. Table 13, p. 31.
6. Wallis, W. A. Compounding probabilities from independent significance tests. *Econometrika*, 1942, **10**, 229-248.

*Manuscript received 10/22/51*

*Revised manuscript received 12/23/51*

# ORTHOGONAL AND OBLIQUE SIMPLE STRUCTURES

W. A. GIBSON

UNIVERSITY OF ILLINOIS

A quick approximation of the best-fitting orthogonal simple structure from the known oblique simple structure is here developed.

## Introduction

For some time the proponents of oblique simple structure in multiple-factor analysis have more or less informally used the cosines of the angles between the oblique reference vectors as ready estimates of the correlations between the corresponding primary factors (cf. 3, p. 212). This is because of the fortunate circumstance that the cells in the matrix showing the cosines of the angles between pairs of reference vectors ( $C = \Lambda'\Lambda$ , where  $\Lambda$  is the oblique transformation matrix) have the same pattern of magnitudes, but with opposite signs in the side entries, as do the cells in the matrix of correlations between the primaries ( $R_{pq} = TT'$ , where each row of  $T$  shows the projections of the associated primary on the initial orthogonal reference frame). An example of this is Thurstone's famous box problem (3, 131-146). The matrices  $C$  and  $R_{pq}$  for that three-dimensional example are shown in Table 1.

TABLE 1\*  
Box Problem  $C$  and  $R_{pq}$

$C = \Lambda' \Lambda$				$R_{pq} = T' T'$			
	$\Lambda_x$	$\Lambda_y$	$\Lambda_z$		$T_x$	$T_y$	$T_z$
$\Lambda_x$	1.000	-.213	-.055	$T_x$	1.000	.229	.105
$\Lambda_y$	-.213	1.000	-.206	$T_y$	.229	1.000	.224
$\Lambda_z$	-.055	-.206	1.000	$T_z$	.105	.224	1.000

\*Reproduced from p. 136 of (3) by permission of the University of Chicago Press.

The purpose of the present paper is to show that this condition can be utilized to provide, for those who are addicted to orthogonal solutions (cf. 2, 42, and 4, 241), a close approximation to the best-fitting orthogonal simple structure when the oblique simple structure position is known. In Figure 1 we have pictured the right spherical triangle corresponding to the best-fitting

orthogonal simple structure for a configuration of test vectors in which the termini of the primaries of the oblique solution are the three points within the triangle and the termini of the reference vectors of the oblique solution are the three points outside the triangle. The appearance of this drawing

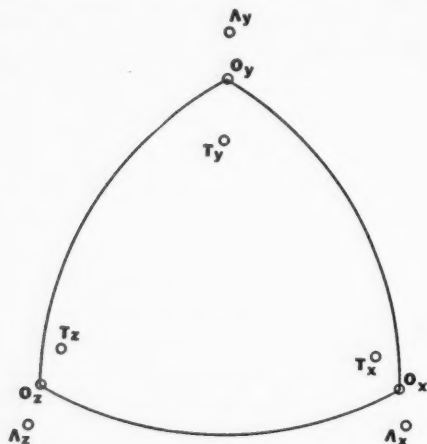


FIGURE 1

suggests that each reference vector of the orthogonal solution is closely approximated by the simple average of the like-numbered primary and reference vectors of the oblique solution. Thus, for example, the unit vector which passes through the centroid of the vectors  $T_x$  and  $\Lambda_x$  ought to be nearly identical with the orthogonal reference vector  $O_x$ .

#### Procedure

To determine the extent to which such a set of centroids will constitute an orthogonal reference frame, we first form the matrix  $\Lambda + T'$  and normalize its columns. The columns of the resulting matrix will give the direction cosines of the centroids of pairs of corresponding primary and reference vectors of the oblique solution. We may designate this matrix as

$$Q = (\Lambda + T')K, \quad (1)$$

where  $K$  is the diagonal matrix of normalizing constants. In Table 2 we have the matrices  $\Lambda$  and  $T'$  for the box problem (3, 136). Also shown in Table 2 are the matrices  $\Lambda + T'$ ,  $K$ , and  $Q$ , along with the necessary computational operations and summational checks. We observe that the non-vanishing entries in  $K$  are very nearly of the same size, which is a trifle greater than .5. This is because each primary and its corresponding reference vector in the box problem are fairly close together. As associated primary and reference

TABLE 2\*  
Box Problem  $\Lambda$ ,  $T'$ ,  $\Lambda + T'$ ,  $K$ , and  $Q$

$\Lambda$				$T'$			
	$\Lambda_x$	$\Lambda_y$	$\Lambda_z$		$T'_x$	$T'_y$	$T'_z$
I	.483	.466	.479	I	.661	.713	.654
II	-.834	.254	.560	II	-.736	.199	.545
III	.267	-.847	.675	III	.144	-.671	.524
$\Sigma$	-.084	-.127	1.714	$\Sigma$	.069	.241	1.723
	$\Lambda + T'$			$Q = (\Lambda + T')K$			
	$x$	$y$	$z$		$Q_x$	$Q_y$	$Q_z$
I	1.144	1.179	1.133	I	.576	.597	.571
II	-1.570	.453	1.105	II	-.791	.229	.557
III	.411	-1.518	1.199	III	.207	-.769	.604
Ch.	-.015	.114	3.437	Ch.	-.008	.058	1.731
$\Sigma$	-.015	.114	3.437	$\Sigma$	-.008	.057	1.732
$\Sigma^2$	3.9426	3.8996	3.9423				
$\sqrt{\Sigma^2}$	1.9856	1.9747	1.9855				
$K$	.5036	.5064	.5037				

\*The matrices  $\Lambda$  and  $T'$  reproduced from p. 136 of (3) by permission of the University of Chicago Press.

vectors approach each other, the corresponding diagonal entry in  $K$  approaches the value .5.

#### *Empirical and Theoretical Verification*

The matrix product  $Q'Q$  will give the cosines of the angles between the unit reference vectors defined by the columns of  $Q$ , and hence will indicate how nearly orthogonal are these reference vectors. The product  $Q'Q$  for the box problem is shown in Table 3. We find that the side entries in Table 3 very nearly vanish, from which we conclude that  $Q$  is quite close to being an orthogonal transformation matrix.

TABLE 3  
Box Problem  $Q'Q$

	$Q_x$	$Q_y$	$Q_z$
$Q_x$	1.000	.004	.013
$Q_y$	.004	1.000	.004
$Q_z$	.013	.004	1.001

Let us now look into the reason for this by carrying out the multiplication  $Q'Q$  symbolically. For this purpose we shall need the relationship

$$T\Lambda = D, \quad (2)$$

which is given by Thurstone (3, 137). Here the diagonal matrix  $D$  shows the cosines of the angles between primary vectors and reference vectors.  $D$  is diagonal because each primary is orthogonal to all but one of the reference vectors and each reference vector is orthogonal to all but one of the primaries. Now we may carry out the multiplication  $Q'Q$  as follows:

$$\begin{aligned} Q'Q &= K'(\Lambda + T')'(\Lambda + T')K \\ &= K(\Lambda' + T)(\Lambda + T')K \\ &= K(\Lambda'\Lambda + \Lambda'T' + T\Lambda + TT')K \\ &= K(C + D' + D + R_{pq})K \\ &= K(C + 2D + R_{pq})K. \end{aligned} \quad (3)$$

The quantity within parentheses in the last step of equation (3) is the sum of four symmetric matrices, two of which are diagonal, while the other two, as we have seen, have side entries which are nearly equal but are opposite in

TABLE 4  
Box Problem  $V_Q$

	$Q_1$	$Q_2$	$Q_3$
1	.990	.119	.050
2	.139	.979	.118
3	.062	.135	.981
4	.574	.812	.112
5	.442	.149	.884
6	.102	.501	.858
7	.795	.614	.075
8	.892	.163	.440
9	.105	.839	.543
10	.703	.719	.099
11	.719	.162	.689
12	.101	.705	.714
13	.980	.148	.043
14	.090	.960	.173
15	.077	.092	.956
16	.370	.491	.777
17	.728	.586	.368
18	.964	.081	.056
19	.181	.945	.071
20	.060	.144	.972

sign. Hence the side entries of this matrix sum are quite small, and the pre- and post-multiplication of this sum by  $K$  makes the resulting side entries even smaller. Since the diagonal entries in  $K$  are all just a little greater than  $1/2$ , we may reasonably summarize the effect of  $K$  by saying that it reduces all the entries in  $C + 2D + R_{pq}$  to a little over  $1/4$  of their original size. Thus any side entry in  $Q'Q$  will be only a trifle larger than  $1/4$  of the algebraic sum of the corresponding entries in  $C$  and  $R_{pq}$ .

From the last step of equation (3) we see that the general diagonal entry in  $K$  is actually defined as the reciprocal of the square root of the corresponding diagonal value in the sum  $C + 2D + R_{pq}$ . That is,

$$k_p = \frac{1}{\sqrt{1 + 2d_p + 1}} = \frac{1}{2} \sqrt{\frac{2}{1 + d_p}}. \quad (4)$$

Thus when  $d_p$  is reasonably large,  $k_p$  is quite close to  $1/2$ . We have already seen this to be true in the box problem.

Applying the transformation matrix  $Q$  to the initial orthogonal factor matrix  $F_0$  by means of the equation

$$F_0 Q = V_Q, \quad (5)$$

we obtain, in  $V_Q$ , the projections of the test vectors on the nearly orthogonal axes defined by the columns of  $Q$ . The matrix  $V_Q$  for the box problem is shown in Table 4. Figures 2, 3, and 4 are the plots of pairs of columns of Table 4. That this solution closely approximates the best-fitting orthogonal simple structure is evident.

### Conclusions

This approximation procedure will work best with configurations whose oblique solutions are not too strongly oblique and are reasonably symmetric with respect to the angles between primaries or reference vectors. In any event, interested persons can find out in a rough way beforehand how good an approximation they will get by comparing the side entries in  $C$  and  $R_{pq}$ . A more precise check on the orthogonality of  $Q$  would be its pre- and post-multiplication by its transpose, and that orthogonality could of course be improved by judicious successive adjustments of the entries in  $Q$ . Additional small rotations of the resulting  $V_Q$  could be made as desired.

After this paper was submitted for publication there was called to my attention an article by Green (1) which presents three different least-squares solutions to this same problem. The procedure described here may be thought of as a rapid approximation to either of Green's two simpler solutions, in which the best orthogonal fit of the reference vectors or of the primary axes is sought, rather than a least-squares minimization of differences in factor loadings as in his first solution.

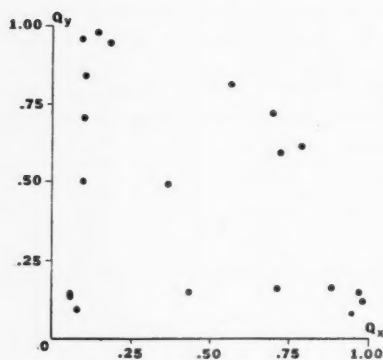


FIGURE 2

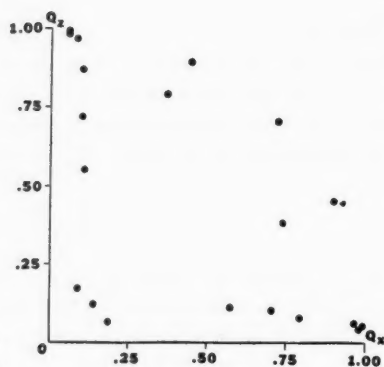


FIGURE 3

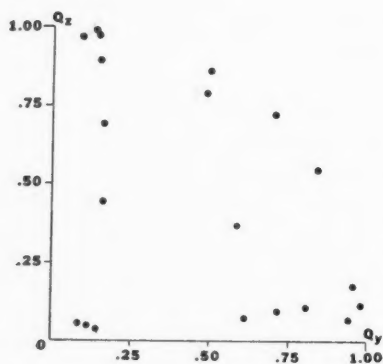


FIGURE 4

*Summary*

We have suggested that the transformation matrix for an orthogonal simple structure solution may often be closely approximated simply by normalizing the columns of the sum of the matrices  $\Lambda$  and  $T'$  of the oblique simple structure solution.

## REFERENCES

1. Green, Bert F., Jr. The orthogonal approximation of an oblique structure in factor analysis. Research bulletin 51-18. Princeton: Educational Testing Service, 1951.
2. Guilford, J. P. Printed classification tests. Report No. 5, Army Air Forces Aviation Psychology Program, Washington: U. S. Government Printing Office, 1947.
3. Thurstone, L. L. Multiple-factor analysis. Chicago: University of Chicago Press, 1947.
4. Wherry, R. J. A new iterative method for correcting erroneous communality estimates in factor analysis. *Psychometrika*, 1949, **14**, 231-241.

*Manuscript received 1/30/52*

*Revised manuscript received 3/26/52*



## COMPANION NOMOGRAPHS FOR TESTING THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN UNCORRELATED PERCENTAGES\*

VALENTINE APPEL

RICHARDSON, BELLOWS, HENRY, & CO.

The purpose of this article is to present a set of nomographs for testing the significance of the difference between uncorrelated percentages, and for determining necessary sample sizes when planning studies which will involve the comparison of percentages. Examples of these two uses are given as well as a description of the formula upon which the nomographs are based.

### *Introduction*

Many short-cut methods of item analysis have been developed in recent years which test the significance of the difference between the proportions of cases in two independent samples responding to a test item. But since most of these techniques have been developed for specific purposes, they tend to suffer limitations in terms of the amount of computation necessary, the precision of the results, or the special assumptions involved particularly concerning sample sizes.

A set of companion nomographs is reproduced here which permits an evaluation of the significance of the difference between uncorrelated percentages while avoiding most of the above limitations.

### *Testing the significance of the difference between uncorrelated percentages*

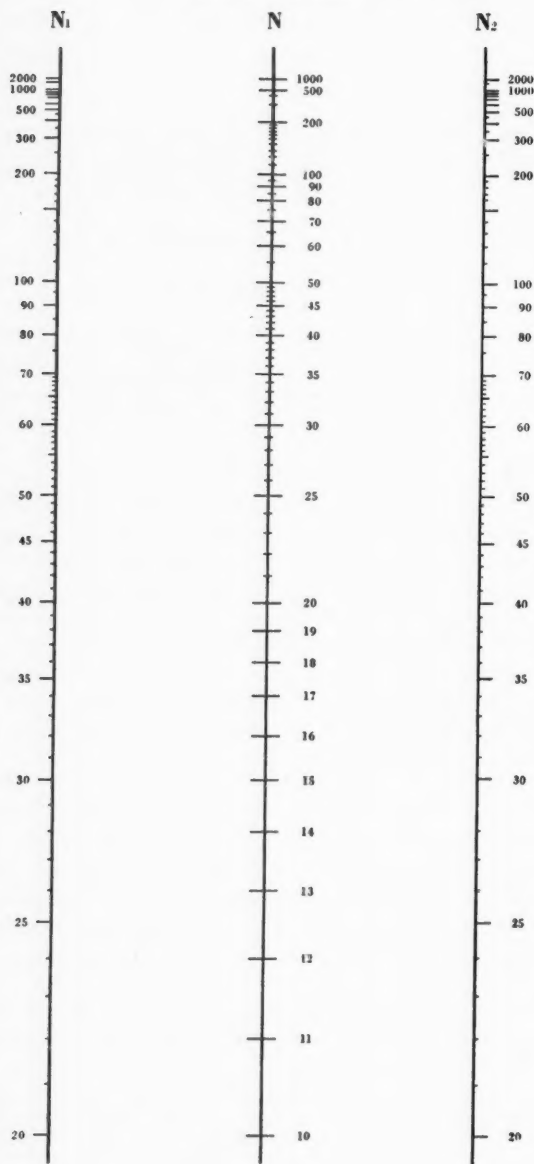
1. Compute the percentage in each of the two samples being compared which respond in a particular way to the item being analyzed. Compute the difference between these two percentages ( $P_1 - P_2$ ). Then compute the percentage responding in this way ( $P$ ) for all available samples combined. In item analysis this might include the middle group as well as the two tail groups.†

2. Enter Nomograph I, connecting the values of  $N_1$  and  $N_2$ , the sizes of the two samples, with a straightedge. Read the value of  $N\ddagger$  from the center scale.

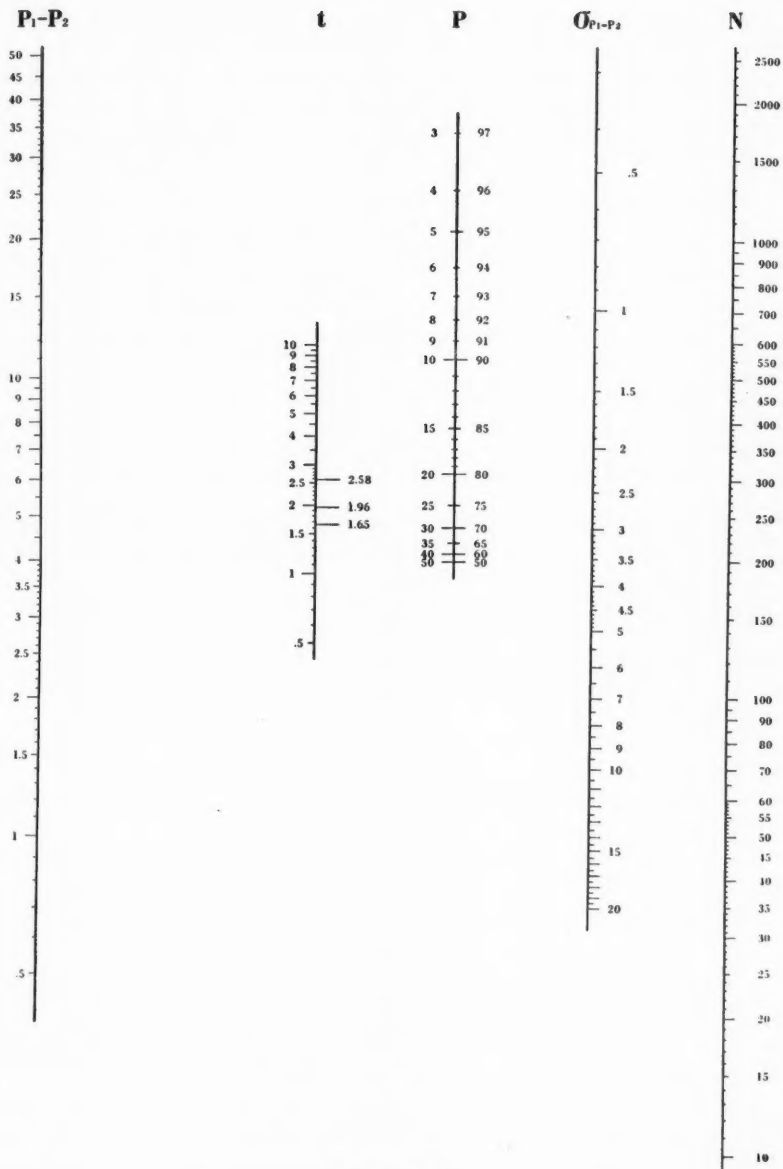
\*The writer wishes to express his thanks and appreciation to Dr. Harold A. Edgerton for his generous advice and assistance in the preparation of this paper, and to Mr. Herman Greenblatt for his painstaking work in preparing these nomographs for the printer.

†Actually, for many practical purposes, a sufficiently stable estimate of  $P$  can be obtained when the middle group is omitted.

‡ $N$ , as employed here, is defined in formula (2) of the section dealing with the derivation of the nomographs.



Nomograph I



Nomograph II

3. Enter Nomograph II, connecting the values of  $N$  and  $P$  with a straight-edge. Read the value of  $\sigma_{P_1-P_2}$  on the center scale.

4. Connect the value of  $\sigma_{P_1-P_2}$  and the value of  $P_1 - P_2$  with a straight-edge. Read the value of  $t$  on the center scale.

Consider the example presented in Table 1. The problem here is to determine whether alternative  $x$  of item  $y$  is able to discriminate the high

TABLE 1  
Per Cent of Cases Marking Alternative  $x$  of Item  $y$

	High Criterion Group	Middle Criterion Group	Low Criterion Group	All Groups Combined
Alternative $x$	59% ( $P_1$ )	49%	42% ( $P_2$ )	49% ( $P$ )
All other alternatives	41%	51%	58%	51%
Sample size	135 ( $N_1$ )	226	163 ( $N_2$ )	524

criterion group from the low criterion group. This can be done by testing the significance of the difference between the percentages marking alternative  $x$  in the high and low criterion groups:

1. The values of  $P_1$  and  $P_2$  have already been computed and are presented in Table 1. The value of  $P_1 - P_2 = 59\% - 42\% = 17\%$ . The value of  $P$  (49%) is shown in the last column of Table 1.

2. Entering Nomograph I, connecting the values of  $N_1$  and  $N_2$ , we find  $N$  to be 74.

3. Entering Nomograph II, connecting the values of  $N$  and  $P$ , we find  $\sigma_{P_1-P_2}$  to be 5.8.

4. Connecting the value of  $\sigma_{P_1-P_2}$  and the value of  $P_1 - P_2$  we find  $t$  to be 2.9.

This  $t$ , when interpreted as a unit normal deviate, is significant beyond the 1% level of confidence. We may conclude, therefore, that alternative  $x$  of item  $y$  is statistically discriminative of these high and low criterion groups.

#### *Determining necessary sample sizes when planning studies*

Although these nomographs have been constructed primarily for the purpose of testing the significance of the difference between percentages, they are also useful in determining necessary sample sizes when planning studies which will involve the comparison of percentages.\*

It frequently happens that investigators decide, before the data collection is begun, that some particular percentage difference shall be considered meaningful. The question then arises as to what sample sizes will be necessary

\*For a previous treatment of this problem see: Swineford, F. Graphical and tabular aids for determining sample size when planning experiments which involve comparisons of percentages. *Psychometrika*, 1946, 11, 43-49.

for this to be true. This problem can be solved simply by referring to Nomographs I and II:

1. Decide upon the percentage difference which will be considered meaningful. Locate this value on the  $P_1 - P_2$  scale in Nomograph II.

2. Decide upon the level of significance required and locate the corresponding  $t$  value.

3. Connect the values of  $P_1 - P_2$  and  $t$  with a straightedge. Read the value of  $\sigma_{P_1-P_2}$ .

4. Connect the value of  $\sigma_{P_1-P_2}$  with the value of  $P$ .<sup>\*</sup> Read the value of  $N$ .

5. Entering Nomograph I with the value of  $N$ , the combinations of permissible sample sizes can be read from the  $N_1$  and  $N_2$  scales by rotating a straightedge around the value of  $N$ . If it is decided that  $N_1$  should equal  $N_2$ , then each sample size will equal  $2N$ .

It will be noted that unless the ratio of the sample sizes ( $N_1/N_2$ ) is known beforehand, as in the case of stratified random sampling, or unless one of the sample sizes is already specified, as would be the case if only a limited sample of a particular type were available, no *unique* solution for determining the necessary sample sizes is possible.

Let us assume that we are about to undertake a study, and that we are willing to accept as meaningful a 10% difference significant at the 5% level. Let us further assume that we have no knowledge of how the combined groups will respond to the items to be employed, and that the sample sizes for each group will be equal:

1. Locate the 10% point on the  $P_1 - P_2$  scale of Nomograph II.

2. Connect this point and the  $t$  value of 1.96 (which is the normal deviate corresponding to the 5% level of confidence) with a straightedge. Reading the value of  $\sigma_{P_1-P_2}$ , we find it to be 5.2.

3. Connect the 50% point on the  $P$  scale with the obtained value of  $\sigma_{P_1-P_2}$ . Reading the value of  $N$ , we find it to be 95.

4. Since we have assumed the two sample sizes to be equal, we multiply  $N$  by 2 in order to determine the two sample sizes.  $N_1$  and  $N_2$  are, therefore, each equal to 190. In other words each sample must contain at least 190 cases before we can be sure that a 10% difference will be significant at the 5% level.

#### Derivation

The formula employed here for the standard error of the difference between two uncorrelated percentages is†

$$\sigma_{P_1-P_2} = \sqrt{PQ\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}, \quad (1)$$

\*If the value of  $P$  (the percentage of cases responding in a given way in the combined sample) is unknown, as it usually is, let  $P$  equal 50%. This will maximize the estimated necessary sample size, thus assuring that percentage differences of the predesignated magnitude will be significant at the required level.

†McNemar, Q. *Psychological Statistics*. New York: John Wiley & Co., 1949.

when  $P$  refers to the percentage of cases in all the combined available samples responding in a particular way to the item being analyzed,  $Q$  equals  $100\% - P$ , and  $N_1$  and  $N_2$  refer to the sizes of the two samples being compared.

Since the technique presented here seeks to test the null hypothesis that all the sample percentages can be considered random samples from the same universe, the most stable estimate of that universe value ( $P$ ) is based upon all the samples combined. Thus in the example presented in Table 1, although the percentage difference being tested was that between the high and the low criterion groups, the best estimate of the universe percentage was based upon the combined sample of high, middle, and low criterion groups.

By definition, let

$$\frac{1}{N_1} + \frac{1}{N_2} = \frac{1}{N}; \quad (2)$$

then formula (1) may be rewritten as follows:

$$\sigma_{P_1-P_2} = \sqrt{\frac{PQ}{N}}. \quad (3)$$

The formula for testing the significance of the difference between two uncorrelated percentages therefore becomes:

$$t = \frac{P_1 - P_2}{\sqrt{\frac{PQ}{N}}}. \quad (4)$$

The functions of Nomographs I and II\* are to solve formulas (2) and (4), respectively.

\*The nomographs reproduced here are too small for efficient use. Single copies of useable size may be obtained free of charge by writing to the author at Richardson, Bellows, Henry & Co., 1 West 57th Street, New York 19, N. Y.

*Manuscript received 1/3/52*

*Revised manuscript received 2/20/52*

## ON THE ECONOMIES OF A PRE-SCREENING TECHNIQUE FOR APTITUDE TEST BATTERIES\*

A. G. ARBOUS AND H. S. SICHEL

NATIONAL INSTITUTE FOR PERSONNEL RESEARCH  
SOUTH AFRICAN COUNCIL FOR SCIENTIFIC AND INDUSTRIAL RESEARCH

Test battery selection is justified on the grounds that it is more efficient than the conventional methods. More precise definitions of efficiency have shown that the additional cost of test administration can often outweigh the gains of selecting a more competent group of persons. Final evaluation of the profit and loss account can be made only when no further economies can be effected in the selection procedures themselves. A psychometric pre-screening technique which will reduce this cost considerably is developed here. Its psychometric structure is mathematically defined, and its advantages illustrated in terms of testing economy (E) and losses (L) of good material for different selection ratios (S. R.). It is shown that, under many circumstances, the testing economy can be considerable, and that in all practical cases "Losses" will be  $\leq 1.5\%$  for all values of  $\rho \geq .8$  (between the pre-screening test and the final battery). The technique in practice is illustrated from a set of data.

### I. Introduction

Previous investigators have devised several techniques for demonstrating the gains of aptitude test procedures over the more conventional methods of selection (1 through 14). One of these authors, Brogden (11, 14) has emphasized that the cost of test administration per hired employee must be taken into consideration before the profit and loss account of testing programmes can be drawn up. He also indicates in these terms how testing can be an unprofitable undertaking in many circumstances.

Even if one concedes that evaluation is to be made solely in terms of a "dollar criterion" and immediate financial gains, Brogden's comments would apply only when the cost of testing cannot be further reduced by some economic pre-screening device—it being remembered that such a device should not result in other losses of potential successes which offset gains.

A psychometric pre-screening device which has these advantages has already been devised by Arbous (8) to meet the need for further economy in a testing program where administration costs were high. In this particular case 56% of the applicant group could be screened out in the first instance by means of a simple pencil-and-paper test, thereby reducing the

\*The authors wish to thank the staff members of the Statistical and Industrial Research Teams of the National Institute for Personnel Research for their assistance in the computational work, preparation of diagrams, and editing of this paper.

application of expensive selection procedures to only 44%. This was achieved at the loss of only 1.4% of those who would have qualified for selection had they completed the full test battery. Gains of this nature are considerable.

It is the object, therefore, of this paper to illustrate, streamline, and put into more general form this pre-screening technique.

## II. The Objectives of Pre-Screening

These have been previously (8) defined as follows:

"(i) to reduce the cost of testing by eliminating in the first instance those candidates who have little or no chance of obtaining the qualifying mark for selection ( $\alpha$ ).

To define quite specifically the operative words "little or no chance," we shall fix the basic condition under which pre-screening is permissible as follows:

The cut-off score ( $\gamma$ ) on the Pre-Screening Test shall be set in such a way that at that level the candidate has a probability of .01 of obtaining or surpassing the Final Battery Score ( $\alpha$ ) at which management has decided selection will ultimately take place. This condition can be more concisely stated as,

$$p[y \geq \alpha | \gamma] = .01. \quad (1)$$

This decision would mean in effect that the only candidates who are screened out are those whose chances are less than 1 in 100 of making the pre-determined battery score ( $\alpha$ ) when fully tested. It will be demonstrated that this safeguard against discrimination is adequate.

One further point must be considered in regard to this basic condition. We have spoken of the battery score at which management has decided to make the final selection ( $\alpha$ ). This is neither a constant level for all batteries, nor for the same battery on all occasions. It will vary according to selection policy (Selection Ratio). This in turn will depend upon the balance to be maintained between manning requirements, on the one hand, and training wastage, on the other. The point is that for any given test battery these conditions of selection will be considered in advance, and the value of  $\alpha$  will be known.

(ii) at the same time pre-screening should ensure that final selection will be no less efficient than it would have been had all applicants passed through the full testing procedures.

Thus, from management's point of view we should ensure that no good material is being lost in the pre-screening process; and from the candidate's point of view, that no individual is discriminated against by being rejected on the basis of one short test, and so prevented from improving his assessment by means of a full testing programme.

Clearly perfection can never be achieved in the assessment of human material, but it will be shown that our basic operative condition in screening,  $p[y \geq \alpha | \gamma] = .01$ , allows us to approach very closely to these desired objectives."

## III. The Theoretical Model

Once the above pre-screening policy has been determined, the development of the technique becomes a matter of finding some predictor ( $x$ ) which will correlate highly with the final battery standard score ( $y$ ). This is usually one of the tests contained in the battery itself. It should preferably be a pencil-and-paper test which is easy to administer on a group basis. Having been included in the final battery score by means of the multiple regression equation, it should have high correlation with the latter. If these conditions can be satisfied, the psychometric structure of the pre-screening test vis-à-vis the final battery score can then be developed in the same way as has been done for the final battery score against the criterion. Operating characteristics can then be drawn up for the pre-screening test as have been done previously for the test battery. The underlying principles and mathematical models are

identical and have already been illustrated (2, 7, 8, and 9). These are immediately applicable to our problem provided the necessary changes are made in terms: viz., the point of dichotomy on the predicted variable is now called  $\alpha$  instead of  $\beta$ , and the new cut-off point on the predictor variable is  $\gamma$  instead of  $\alpha$ .

We may thus write Sichel's (7) formula for the selector's operating characteristic as

$$p[y \geq \alpha | x] = \Phi \left[ \frac{\alpha - \rho x}{\sqrt{1 - \rho^2}} \right], \quad (2)$$

where

$x$  = pre-screening test scores in standard measures,

$y$  = battery test scores in standard measures,

$p[y \geq \alpha | x]$  = the conditional probability of a pre-screening score ( $x$ ) being associated with a score on the battery ( $y$ ) which is  $\geq$  a pre-determined value  $\alpha$ ,

and

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(z^2/2)} dz.$$

By integrating the product of equation (2) and the normal distribution

$$\text{erf}(x) = \frac{1}{\sqrt{2\pi}} e^{-(x^2/2)},$$

we get the proportion of the applicant population who are both selected by the pre-screening test and make or surpass the battery score  $\alpha$ , i.e.,

$$p_1 = \int_{\gamma}^{+\infty} \Phi \left[ \frac{\alpha - \rho x}{\sqrt{1 - \rho^2}} \right] \text{erf}(x) dx. \quad (3)$$

From Figure 1 we see that the proportion of applicants who are not selected by the pre-screening test but equal or exceed the battery cutting score  $\alpha$  is

$$p_4 = \Phi(\alpha) - p_1. \quad (4)$$

Hence it follows that the proportionate loss of good material for a given set of pre-screening conditions is given by

$$\begin{aligned} \psi_{1,\gamma} &= \frac{\text{Number of candidates having battery scores } \geq \alpha \text{ lost in pre-screening}}{\text{Total number of candidates in population having battery scores } \geq \alpha} \\ &= \frac{p_4}{p_1 + p_4} = 1 - \frac{p_1}{\Phi(\alpha)} = \frac{1}{\Phi(\alpha)} \int_{-\infty}^{\gamma} \Phi \left[ \frac{\alpha - \rho x}{\sqrt{1 - \rho^2}} \right] \text{erf}(x) dx. \end{aligned} \quad (5)$$

We now have the necessary information to construct the mathematical model for the pre-screening technique and illustrate how it works in the case of a specific test correlating, e.g., .80, .85, or .90 with the full battery.

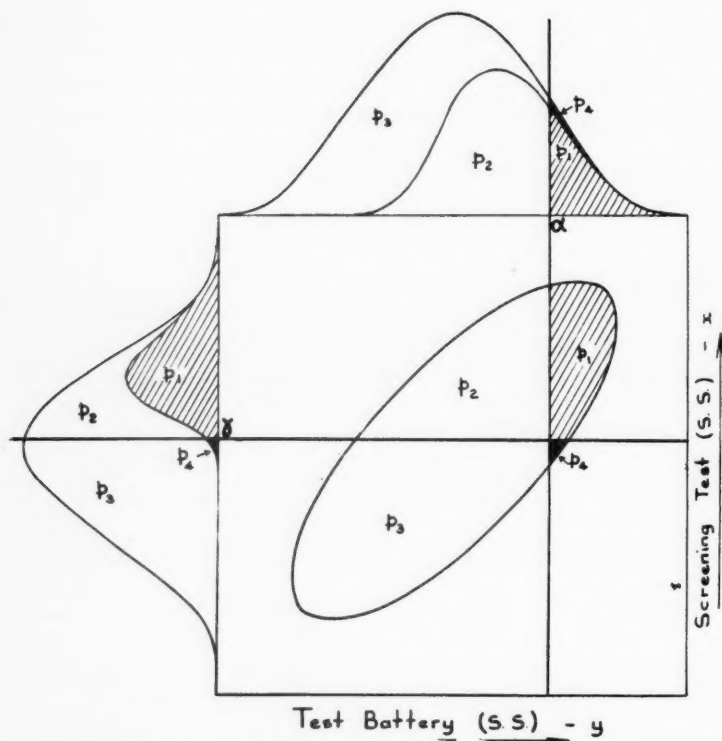


FIGURE 1

Testing Economy ( $E$ ) and Loss ( $L$ ) Achieved by Pre-screening Technique

The results will be reported in terms of the following:

- (1) The value of  $\gamma$  for a given selection ratio and a probability of success of .01

Without pre-screening, the selection ratio (S.R.) pertaining to the full battery is given as  $\Phi(\alpha)$ . With pre-screening, there is a certain loss  $\psi_{1,\gamma}$  defined in (5) which reduces the actual S.R. obtained after the application of this technique. To accommodate this loss, the selection ratio resulting after pre-screening is precisely defined as

$$\text{S.R.} = (1 - \psi_{1,\gamma})\Phi(\alpha). \quad (6)$$

Equation (6) requires *a priori* knowledge of  $\alpha$  and  $\gamma$ . In practice, however, it is the selection ratio which is the given quantity. We have to find  $\alpha$  and  $\gamma$  subject to

$$p[y \geq \alpha \mid \gamma] = .01$$

as previously laid down. For this condition we have from equation (2)

$$\Phi \left[ \frac{\alpha - \rho \cdot \gamma_{.01}}{\sqrt{1 - \rho^2}} \right] = .01$$

$$= \Phi[2.32635].$$

Hence,

$$\gamma_{.01} = \frac{\alpha - 2.32635 \sqrt{1 - \rho^2}}{\rho}. \quad (7)$$

The solution of equations (6) and (7) for the required  $\alpha$  and  $\gamma_{.01}$  is not possible by elementary means but may be effected with the help of graphs as shown at a later stage.

(2) *The "testing economy" achieved by the pre-screening technique*

Let us define "testing economy" ( $E$ ) as the percentage of candidates who, with the exception of the pre-screening procedure, need not undergo the majority of tests making up the battery. From Figure 1 we deduce

$$E = 100[1 - \Phi(\gamma_{.01})]\%. \quad (8)$$

Often  $E$  is large enough to cause a major saving in testing costs.

(3) *The percentage of potential material lost in pre-screening*

The percentage of candidates who would make or surpass the battery selection score  $\alpha$  but are lost in the application of the pre-screening technique has already partly been defined in equation (5). For the given condition  $p[y \geq \alpha|\gamma] = .01$ , we write the "loss" as

$$L = 100\psi_{1, \gamma_{.01}}\%. \quad (9)$$

TABLE 1  
 $\gamma_{.01}$ ,  $E$ , and  $L$  for Various Selection Ratios\*  
( $\rho = .90$ )

Selection Ratio S.R. (in %)	Battery cutting Score $\alpha$	Pre-screening cutting score $\gamma_{.01}$	Testing Economy $E$ (in %)	Loss $L$ (in %)
2.23	70	60.96	86.35	1.86
6.62	65	55.40	70.54	0.94
15.80	60	49.84	49.36	0.41
30.80	55	44.29	28.40	0.17
49.97	50	38.73	12.99	0.06
69.13	45	33.18	4.63	0.02

\*In this study raw test scores were converted to standard scores having a mean of 50 and standard deviation of 10.

(4) *The Calculation of  $\gamma_{.01}$ ,  $E$ , and  $L$  for Given Selection Ratios and Correlations*

These quantities were calculated with the help of Arbous' Tables (9) and the above formulas, and are reflected in Table 1 together with the correct pre-screening cutting score  $\gamma_{.01}$ .

From the above table we see that for a selection ratio of 6.6% and a correlation  $\rho = .90$ , the majority of applicants (70.5%) need not be tested on the full battery. None of them has a chance exceeding 1 in 100 to be selected by the total combination of tests. This major saving in testing costs is achieved at a loss of only 0.9% of good material; i.e., for 100 candidates who should all be selected by the full battery, only one would be eliminated by the pre-screening test.

The results embodied in Table 1 are graphed in Figures 2, 3, 4, and 5, taking the selection ratio as the independent variable.

For the benefit of those who wish to prepare tables and graphs suiting their own needs, we have repeated the necessary calculations based on  $\rho = .80$  and  $.85$  as well. The results are shown in Table 2. The corresponding graphs based on these values are included also in Figures 2, 3, 4, and 5.

(5) *Appraisal of Results*

In Figure 2 the three curves coincide on the graph for the three levels of  $\rho = .80$ ,  $.85$  and  $.90$ . This is due to the minute loss which is sustained in pre-screening. In fact, for most practical purposes, one could disregard this loss in deriving values of  $\alpha$ , and use directly the tables for the normal integral.

The most striking feature of Figure 4 is the fact that the differences in testing economy ( $E$ ) between the different values of  $\rho$  are greatest in the middle ranges of S.R. which will occur more often in practice. Here a difference of .1 in  $\rho$  can have a considerable effect on  $E$ . E.g., when S.R. = 20%,

$$\rho = .80, \quad E = 24.0\%$$

$$\rho = .90, \quad E = 42.5\%.$$

In this particular case the testing economy for  $\rho = .90$  is nearly twice that of  $\rho = .80$ .

From Figure 5 it will be noted that for large S.R.'s a difference in  $\rho$  from  $.80$  to  $.90$  makes virtually no difference in the loss of good material. This difference is exceedingly small. It becomes relatively larger, however, as S.R. decreases, but in most practical cases (i.e., when S.R.  $\geq 7\%$ ) the losses will be  $< 1.5\%$  for all values of  $\rho \geq .80$ .

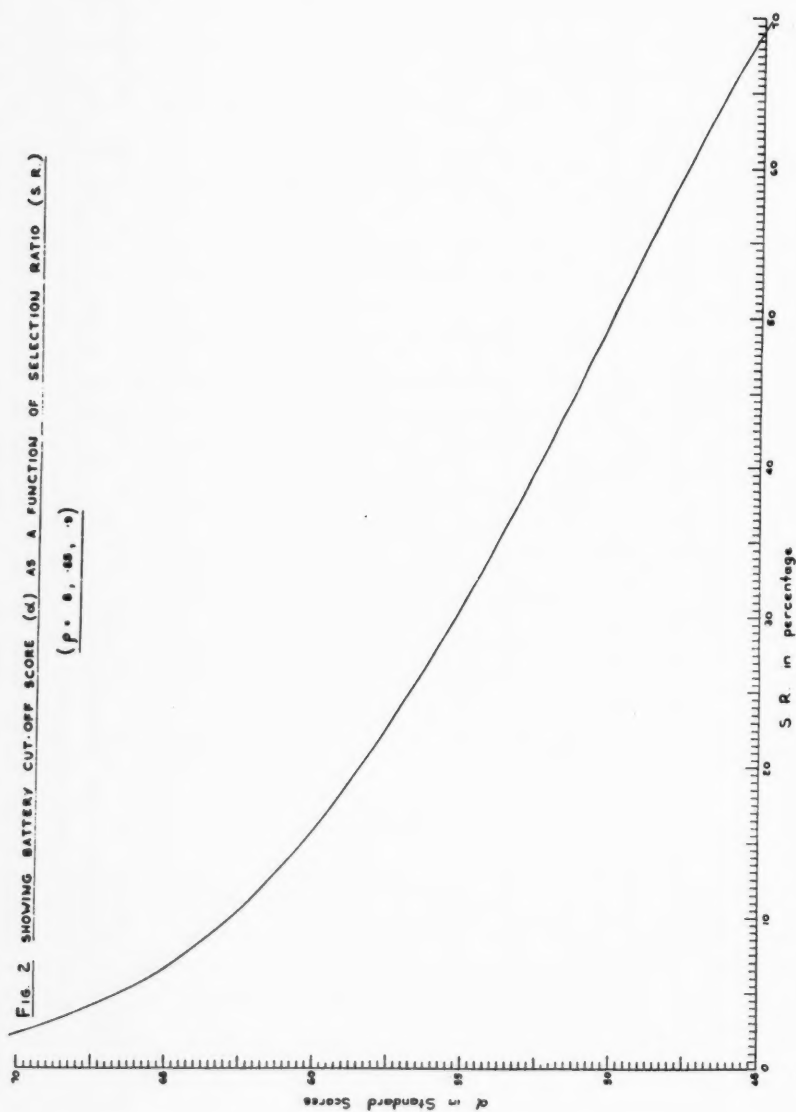
IV. *An Alternative Model*

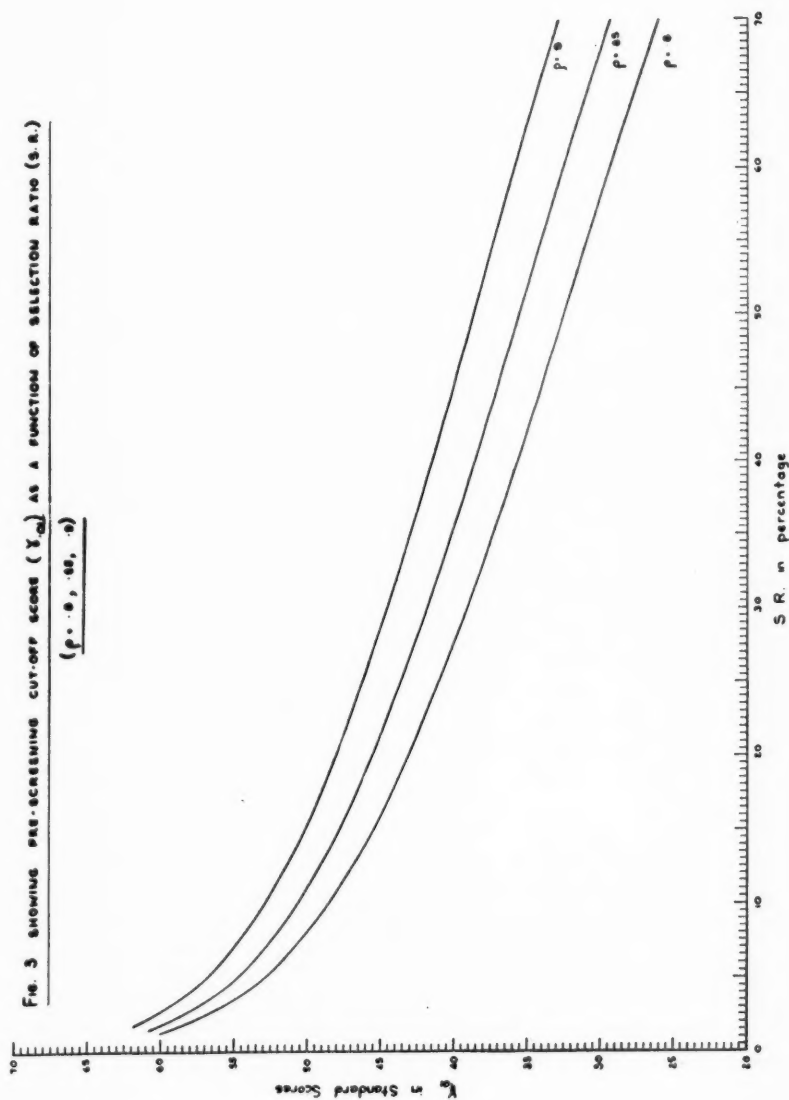
In the example given above we have set the operative condition in pre-screening as  $p[y \geq \alpha|\gamma] = .01$ , and, having done so, indicated to management what losses ( $\psi_{1,\gamma_{.01}}$ ) it suffered at the different levels of S.R.

TABLE 2  
 $\alpha$ ,  $\gamma_{.90}$ ,  $E$ , and  $L$  as Functions of Selection Ratio (S.R.) and Correlation Coefficient ( $\rho$ )\*

$\rho = .80$					$\rho = .85$					$\rho = .90$				
S.R.	$\alpha$	$\gamma_{.90}$	$E$	$L$	S.R.	$\alpha$	$\gamma_{.90}$	$E$	$L$	S.R.	$\alpha$	$\gamma_{.90}$	$E$	$L$
2	70.4	58.1	78.9	4.0	2	70.4	59.6	83.1	2.9	2	70.4	61.3	87.0	1.9
3	68.7	55.9	72.8	3.1	3	68.7	57.6	77.7	2.4	3	68.7	59.5	82.9	1.6
4	67.4	54.3	67.3	2.4	4	67.4	56.1	72.9	2.0	4	67.4	58.1	79.1	1.4
5	66.3	53.0	62.3	2.0	5	66.3	54.9	68.6	1.6	5	66.3	56.9	75.6	1.2
6	65.4	51.9	57.8	1.7	6	65.4	53.9	64.8	1.4	6	65.4	55.9	72.4	1.0
7	64.6	51.0	53.8	1.4	7	64.6	53.0	61.4	1.2	7	64.6	55.0	69.4	.9
8	63.9	50.2	50.2	1.2	8	63.9	52.2	58.3	1.0	8	63.9	54.2	66.6	.8
9	63.3	49.4	47.0	1.1	9	63.3	51.5	55.5	.9	9	63.3	53.5	63.9	.7
10	62.8	48.6	44.1	.9	10	62.8	50.8	52.8	.8	10	62.8	52.8	61.3	.7
15	60.3	45.5	32.6	.5	15	60.3	47.8	41.3	.5	15	60.3	50.2	50.9	.4
20	58.4	43.2	24.0	.3	20	58.4	45.6	32.2	.3	20	58.4	48.1	42.5	.3
25	56.7	41.1	17.9	.2	25	56.7	43.7	25.4	.2	25	56.7	46.3	35.5	.2
30	55.2	39.2	13.7	.1	30	55.2	42.0	20.3	.2	30	55.2	44.6	29.6	.2
35	53.8	37.4	10.5	.1	35	53.8	40.4	16.2	.1	35	53.8	43.0	24.6	.1
40	52.5	35.7	8.0	.1	40	52.5	38.8	12.8	.1	40	52.5	41.5	20.3	.1
45	51.2	34.1	5.9	.1	45	51.2	37.2	9.9	.1	45	51.2	40.1	16.5	.1
50	50.0	32.6	4.1	0	50	50.0	35.6	7.5	.1	50	50.0	38.7	13.0	.1

\*Due to rounding, tabulated values may be in error by one in the last digit.





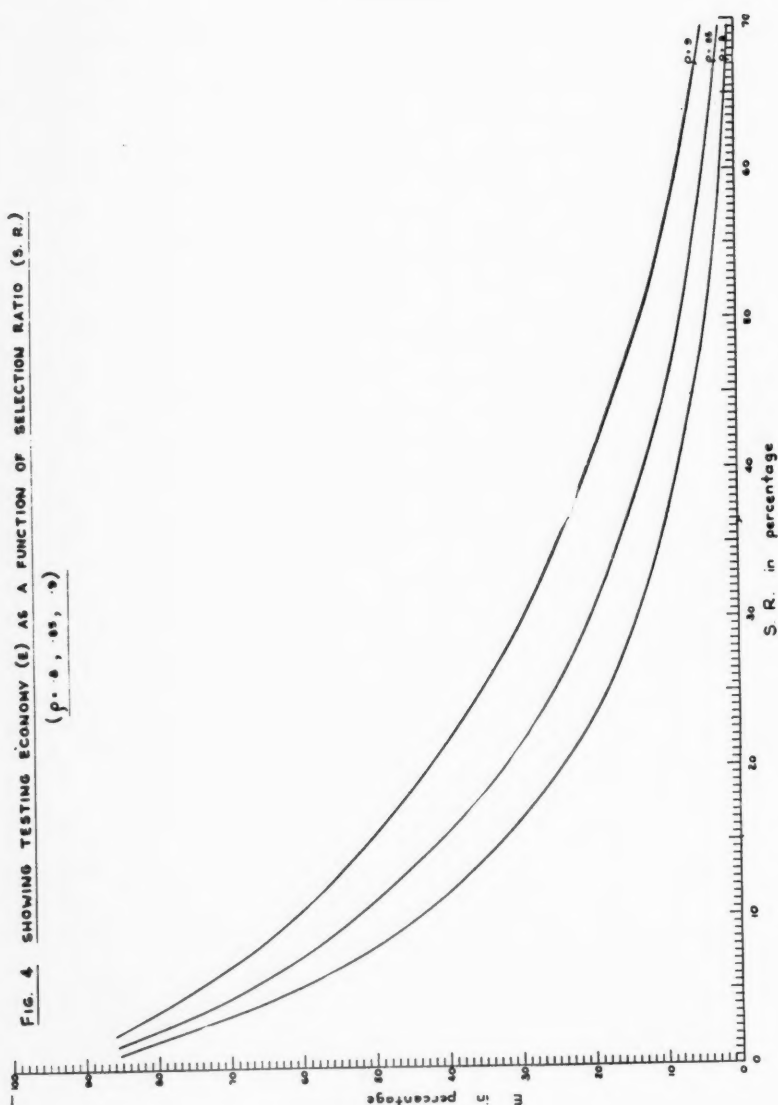
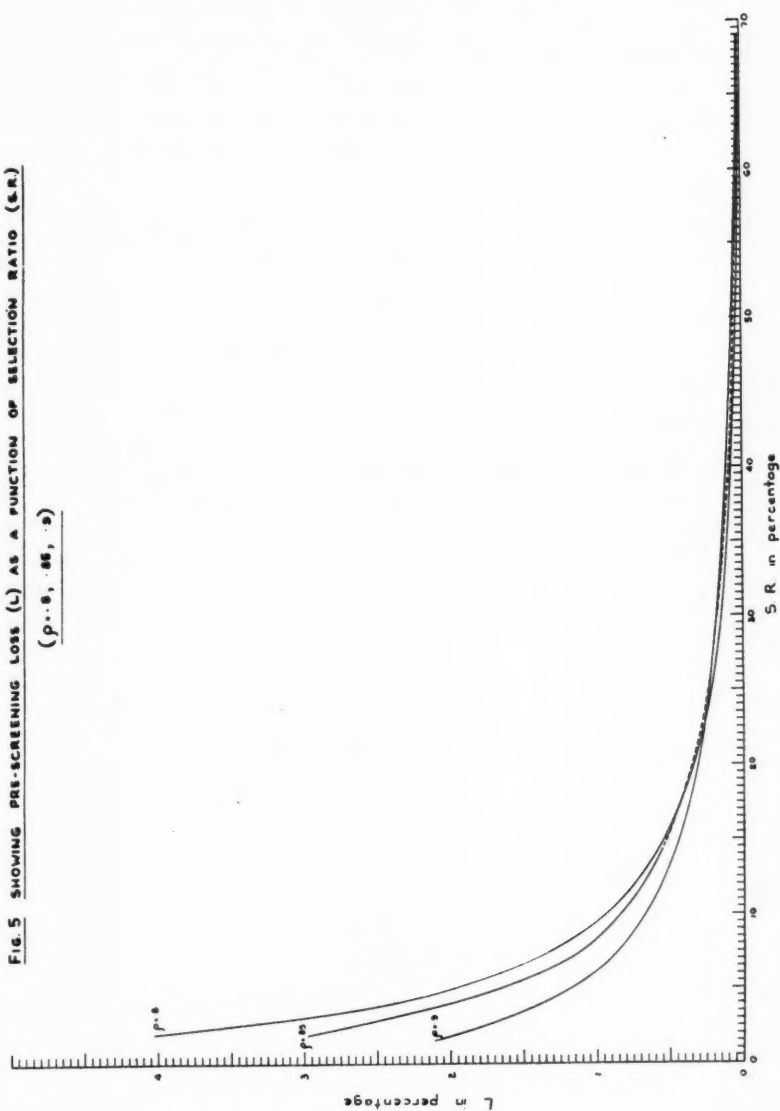


FIG. 5 SHOWING PRE-SCREENING LOSS (L) AS A FUNCTION OF SELECTION RATIO (S.R.)  
 ( $p=0.9, 0.8, 0.7$ )



It would have been just as possible to set the condition in management's interest as  $\psi_{1,\gamma} = .01$ , i.e., to permit a loss of 1% of the candidates who would make or surpass the pre-determined battery selection score  $\alpha$ , under all circumstances. It would merely be necessary then in turn to report to the candidates at what level of probability of making this score they would have been pre-screened out, for different values of S.R., i.e., to determine the resulting value of  $p[y \geq \alpha|\gamma]$ .

This alternative method was fully investigated and was rejected because of the fact that for certain values of  $\alpha$  which were likely to be met in practice, the value of  $p[y \geq \alpha|\gamma]$  became too high. E.g., when

$$\alpha \leq 0.5 \text{ s.m.}, \quad p[y \geq \alpha|\gamma] \geq .04$$

$$\alpha = 0.0 \text{ s.m.}, \quad p[y \geq \alpha|\gamma] = .08$$

in the case of both  $\rho = .80$  and  $.90$ .

#### V. The Technique in Practice

It now becomes necessary to test the mathematical model for the pre-screening technique by comparing the expected results with those actually obtained in a practical situation.

In the selection program for potential army recruits a battery of four pencil-and-paper tests is administered to the applicants. The intercorrelation matrix of the tests based on a sample of 1637 subjects was given in a previous study by Sichel (16) as:

	M	B	H
A(G)	.637	.596	.690
M		.534	.442
B			.444

From

$$R_{i,B} = \frac{\sum_{i=1}^k \beta_i r_{ii}}{\sqrt{\sum_{i=1}^k \sum_{j=1}^k \beta_i \beta_j r_{ij}}}, \quad (10)$$

where

$R_{i,B}$  = correlation between  $i$ th test and total battery (including that test),

$\beta_i$  = battery weight of  $i$ th test, and

$r_{ij}$  = correlation between  $i$ th and  $j$ th test,

we may find the required correlation between the pre-screening test and the full battery. Using A(G) and all  $\beta = 1/4$  we have

$$R_{i,b} = .894.$$

The correlation based on a scattergram of test A(G) and the complete battery scores (both variables in standard scores) checked the above derived correlation to the last decimal.

The formulas given above all presuppose a bivariate normal distribution. To test this assumption a bivariate normal surface was fitted to the scattergram using Pearson's Tables (15). For 36 degrees of freedom and  $\chi^2 = 49.9$ , we found  $P = .064$  so that the hypothesis of binormality cannot be rejected.

For a given set of  $\alpha$ 's advancing from zero in half sigma units corresponding values of  $\gamma_{.01}$  were calculated from equation (7). A new scattergram was drawn up using the matching  $\alpha$  and  $\gamma_{.01}$  values as class boundaries. The scattergram is shown in Figure 6. It enables us to compare theory and observations by counting the number of applicants who are screened out, being lost

**Test Battery in Standard Scores**

	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80	80-85	
77-49										1	4	8	2	4	19
71-90															
71-90							1	1	5	10	26	13	4	1	61
66-30															
66-30								8	33	68	32	9	1		151
60-71															
60-71					1	1	19	81	92	50	12	1			257
55-12															
55-12					3	22	96	145	87	10	3				366
49-53															
49-53					24	85	149	75	15	1					345
43-93															
43-93			1	13	50	115	30	15							224
38-34															
38-34		2	5	24	57	38	9								135
32-75															
32-75		5	7	23	18	4		1							58
27-16															
27-16	1	1	7	4											13
21-56															
21-56		1	3												4
15-97															
	1	9	23	64	153	265	304	326	232	140	77	31	7	5	1637

FIGURE 6  
Bivariate Frequency Distribution of Pre-Screening Test Against Full Battery

TABLE 3  
Comparison of Theoretical and Actual Values of  $E$ ,  $L$ , and S.R.

$\alpha$	$\gamma_{.01}$	Testing Economy ( $E\%$ )		Loss ( $L\%$ )		Selection Ratio (S.R.%)	
		Expected	Observed	Expected	Observed	Expected	Observed
70.0	60.7	85.8	85.9	2.0	2.3	2.2	2.6
65.0	55.1	69.6	70.2	1.0	2.5	6.6	7.3
60.0	49.5	48.1	47.8	0.4	0.4	15.8	15.9
55.0	43.9	27.2	26.5	0.2	0.0	30.8	30.1
50.0	38.3	12.2	12.8	0.1	0.1	50.0	50.0
45.0	32.8	4.2	4.6	0.0	0.1	69.1	68.5

The use of Table 2 is illustrated by the following example:  
Given:  $R_{t,B} = .894$ , S.R. = 15.8%. Find  $\alpha$ ,  $\gamma_{.01}$ ,  $E$ , and  $L$ .

by pre-screening, etc. for matching values of  $\alpha$  and  $\gamma_{.01}$ . The results are summarized in Table 3. The close agreement between theory and observations will be noted.

The detailed linear interpolations in Table 4 based on Table 2 are self-explanatory. The slight discrepancies between the last row of Table 4 and the third row of Table 3 are due to rounding errors in the tabular values used.

The above example may also be used to illustrate the method for obtaining the observed values of  $E$  and  $L$  from the bivariate distribution in Figure 6.

In accordance with Figure 1 heavy lines were drawn in Figure 6 vertically

TABLE 4  
Solution by Interpolation of  $\alpha$ ,  $\gamma_{.01}$ ,  $E$ , and  $L$   
for Illustrative Problem with  $r = .894$  and S.R. = 15.8%

$\rho$	S.R.(%)	$\alpha$	$\gamma_{.01}$	$E\%$	$L\%$
.85	15	60.3	47.8	41.3	0.5
.85	20	58.4	45.6	32.2	0.3
.85	15.8	60.0	47.45	39.84	0.47
.90	15	60.3	50.2	50.9	0.4
.90	20	58.4	48.1	42.5	0.3
.90	15.8	60.0	49.86	49.56	0.38
.894	15.8	60.0	49.6	48.4	0.4

at  $\alpha = 60$ , and horizontally at  $\gamma = 49.526$  for S.R. = 15.8%. The discrepancy between the value for  $\gamma$  used and that obtained in Table 4 is due to the effects of rounding the last decimal in the calculations. The actual value in Table 4 is 49.57. The theoretical value was 49.526 which we used for the bivariate distribution.

From Figure 6 it will be seen that

- (i) in all, 260 persons make a battery score of  $\alpha \geq 60$ . Of these only 1 is lost by being pre-screened out (i.e., having a  $\gamma$  score = 49.526). The loss ( $L$ ) is thus  $1/260 = 0.4\%$ .
- (ii) 783 persons are pre-screened out and need to do only 1 test. The testing economy ( $E$ ) is thus  $783/1637 = 47.8\%$ .

#### VI. The Dollar Value of Pre-Screening Economies

It would have been admirable to express the cost of battery testing and pre-screening as functions of some variable, and by some derived equation, to have graphed the actual dollar savings of the technique for different testing demands and different S.R.'s.

The fact that this is not possible will be readily appreciated by those who have had experience of diverse testing programs.

It is hoped that the model developed here will enable testers to estimate financial savings in their own particular circumstances.

#### REFERENCES

1. Taylor, H. C., and Russell, J. T. The relationship of validity coefficients in the practical effectiveness of tests in selection: tables and discussions. *J. appl. Psychol.*, 1939, 23, 565-578.
2. McClelland, W. Selection for secondary education. Univ. London Press, 1942.
3. Bittner, R. H., and Wilder, C. E. Expectancy tables: a method of interpreting correlation coefficients. *J. exp. Educ.*, 1946, 14, 245-252.
4. Berkson, J. "Cost utility" as a measure of the efficiency of a test. *J. Amer. statist. Ass.*, 1947, 42, 246-255.
5. Jarrett, R. F. Percent increase in output of selected personnel as an index of test efficiency. *J. appl. Psychol.*, 1948, 32, 135-145.
6. Bingham, W. V. Great expectations. *Personnel Psychol.*, 1949, 2, 397-404.
7. Sichel, H. S. The selective efficiency of a test battery. *Psychometrika*, 1952, 17, 1-39.
8. Arbous, A. G. Selection for industrial leadership. (To be published by Oxford Univ. Press).
9. Arbous, A. G. A set of tables and graphs for the operating characteristics of aptitude test batteries. (To be published).
10. Brogden, H. E. On the interpretation of the correlation coefficient as a measure of predictive efficiency. *J. educ. Psychol.*, 1946, 37, 65-76.
11. Brogden, H. E. When testing pays off. *Personnel Psychol.*, 1949, 2, 170-183.
12. Brogden, H. E. An approach to the problem of differential prediction. *Psychometrika*, 1946, 11, 139-154.
13. Brogden, H. E. Increased efficiency of selection resulting from replacement of a single predictor with several differential predictors. *Educ. psychol. Meas.*, 1951, 11, 173-195.

14. Brogden, H. E., and Taylor E. K. The dollar criterion—applying the cost accounting concept to criterion construction. *Personnel Psychol.*, 1950, 3, 133-154.
15. Pearson, K. Tables for statisticians and biometricians, parts I and II. Cambridge: Cambridge Univ. Press, 1931.
16. Sichel, H. S. First peace-time validation of army selection tests with a discussion of some statistical problems encountered in this project. Bulletin of the National Institute for Personnel Research, South Africa, No. 1, Vol. II, January, 1950.

## THE FACTORIAL COMPOSITION AND VALIDITY OF DIFFERENTLY SPEEDED TESTS

CHARLES T. MYERS

EDUCATIONAL TESTING SERVICE

Scores on five differently speeded parts of a nonverbal reasoning test were analyzed to determine the relationship between the abilities measured by the speeded and the unspeeded parts. Two orthogonal factors were found: One is described as the ability to answer the problems correctly; the other is described as the tendency to answer the problems quickly. The first factor was found to be somewhat more valid than the second for predicting grades at the U. S. Naval Academy. The suggestion is made that descriptions of tests are incomplete unless they include some specification describing the speededness of the test.

### *Introduction*

The purpose of this study was to compare the validities of short nonverbal reasoning tests when different numbers of items were given within the same time limits. The criteria for validity were grades and grade averages for the first year at the U. S. Naval Academy at Annapolis. Three forms of a figure-classification test were given in 1948 to 600 of the midshipmen who were then in their first summer at the Academy. The forms were made up of five 12-minute parts, these parts including either 10, 20, or 30 items with at least one part of each form at each speed level. Correlations were computed between scores on the different parts, and also between the parts and the grades in the first year's work at the Academy. These intercorrelations were analyzed by factorial methods. The following conclusions were drawn from this study: that the moderately speeded parts were the most valid and that the ability to answer the problems correctly showed no correlation with the tendency to answer the problems quickly.

### *Description of the Test*

Figure-classification problems of the type used in this study consist of two sets of four geometric figures which define the problem, and of five multiple-choice answer figures. The persons taking the test are required to induce the characteristic possessed by the first four figures but not by the second four and to select the only one of the five answer figures which has the characteristic common to the first four problem figures. Such characteristics may involve the size, shape, position, drawing, or number of parts of the figures.

For this experiment a test of 100 figure-classification items was prepared in three different forms. These items were printed on ten pages with ten items on a page. The same ten pages were used in each form. In all three forms these pages were divided into five 12-minute parts with either one, two, or three pages to a part. The three forms differed only in the grouping of the pages into these parts of various lengths. The parts of Form A had ten, twenty, thirty, ten, and thirty items in that order. The parts of Form B had twenty, thirty, ten, ten, and thirty items. The parts of Form C had thirty, ten, twenty, ten, and thirty items. The order of the pages and their division into parts is shown in Table 1.

TABLE 1  
Order of Pages and  
Speededness of Parts as Indicated by Per Cent Completing Each Page\*

		<i>Form A</i>	<i>Form B</i>	<i>Form C</i>
Part 1:	1st Page	97 (a)†	99 (a)	100 (a)
	2nd Page		70 (b)	79 (b)
	3rd Page			41 (d)
Part 2:	1st Page	98 (c)	99 (c)	91 (c)
	2nd Page	57 (d)	77 (d)	
	3rd Page		37 (f)	
Part 3:	1st Page	99 (e)	100 (e)	100 (e)
	2nd Page	65 (f)		65 (f)
	3rd Page	30 (b)		
Part 4:	1st Page	100 (g)	99 (g)	100 (g)
Part 5:	1st Page	100 (h)	99 (h)	100 (h)
	2nd Page	84 (i)	82 (i)	82 (i)
	3rd Page	36 (j)	34 (j)	38 (j)

\*Per cent of subjects attempting the last item on each page or some subsequent item in the same part.

†Letters in parentheses identify pages containing identical sets of ten items.

This arrangement made it possible to observe whether practice on the test would change the effect of speededness. Since each different speed level was used as part one of some form and since in each form one of the last two parts was unspeeded and the other highly speeded, it was possible to compare each speed level when given without any practice, with the extreme speed levels after 36 minutes and 60 items of practice.

#### *Administration of the Test*

The test was administered to a group of approximately 600 midshipmen during their first summer at the U. S. Naval Academy.\* The three forms of

\*Grateful acknowledgement is made to the officers of the Naval Academy for their assistance in administering the test and for providing the data for the validation of the test.

the test were distributed in rotation according to the order in which the men happened to be seated, each man taking only one form. All three forms were administered simultaneously. The directions on the cover page included the following statement: "Not all the parts have the same number of problems; the number is printed at the beginning of each part." The cover page also indicated that twelve minutes were to be allowed for each part.

#### *Test and Criterion Scores*

Nineteen scores were obtained for each form. These scores were as follows: For each of the ten pages the number of problems correct; for each two-page part the number of problems correct; and for each three-page part the number of problems correct, the number of problems wrong, the number skipped, and the number attempted. By "number skipped" is meant the number of unanswered problems preceding the last problem that was answered; by "number attempted" is meant the number of the last problem that was answered.

After the academic year which followed the administration of the test the Naval Academy furnished course grades of the men for that year in the form of rank-order scores. The men were graded on the following: Conduct, Physical Training, Marine Engineering, Mathematics, Electrical Engineering, English, Foreign Languages, and a Relative Standing which is a weighted average of all the grades. These rank-order scores were converted to normal distributions with means of 50 and standard deviations of 10 for the whole class.

#### *Speededness of the Parts*

The speededness of the parts of the test may be judged by the percentages of persons marking responses at various points through the test. This is also shown in Table 1. The most speeded parts were finished by from 30 to 41 per cent of the subjects. In four-fifths of all the parts the tenth problem was attempted by ninety-nine per cent of the subjects. None of the parts was a "pure speed test" since they all contained some difficult items. For the 30-item parts the variance of the number of items wrong was about equal to the variance of the number of items attempted.

#### *Analysis of the Intercorrelations*

Product-moment correlations between all the test and criterion scores were obtained by I. B. M. machine methods. The three tables of intercorrelations were analyzed by Thurstone's grouping method. This was done to help evaluate the effect of the time limits on the validity of the test. Scores on the first page of each of the five parts were used as a group to define the first factor. Scores on the last page of each of the two three-page parts together with the scores on the number of items attempted on the same two parts were used as a group to define the second factor. The two groupings

were selected because the scores in each group had high intercorrelations among themselves and low intercorrelations with other scores, and because the proportionality of the rows of intercorrelations suggested that their vectors would form clusters. After the extraction of the second factor the small size of the residuals indicated that no more factors could be extracted from the page-score correlations, and the correlations between test scores and the criterion scores were mostly accounted for by these two factors. For the three matrices a distribution was made of these residuals after the second factor had been extracted. This distribution did not include the residuals of the intercorrelations between criterion scores or between the wrong, skip, or number-attempted scores. None of the 546 residuals in this distribution had an absolute value of more than .17 and only 28 of the residuals had an absolute value of more than .10. Over two thirds of the residuals were within the range of .05 to  $-.05$ .

The complete matrices included a few cells expressing relationships between experimentally dependent scores. In these cells values were substituted from other parts of the same matrix. For each cell expressing the relationship of dependent scores there were two other cells, one in the same row and the other in the same column, expressing the relationship of those two scores with similar but independent measures. The substitute value that was used was the arithmetic mean of those two cell entries. Preliminary investigation using five independent variables, three for the unspeeded and two for the speeded parts, indicated that the above procedure would not distort the configuration of vectors.

The two factors which were found in each of the matrices were able to account for practically all the correlation between the test scores as well as for the validity of the tests. Factor loadings are shown in Table 2. The first factor that was extracted was defined by the scores of the number right on the first pages of all the parts, of whatever length, and may be described as the ability to answer the problems correctly. The second factor was defined by either the number of problems answered correctly on the last pages of the 30-problem parts, or the number of problems attempted in those parts. This factor may be described as the tendency to answer the problems quickly.

All three analyses resulted in similar configurations of test and criterion vectors. The fifteen scores used to define the first factor in the three analyses had loadings on that factor in the range from .51 to .75. The use of the grouping method without rotation necessarily resulted in some positive and some negative loadings on the second factor for these vectors in each analysis. The homogeneity of these scores can be judged from the observation that their loadings on the second factor fell within the range .08 to  $-.10$ . In this method the factors must be orthogonal, but the loadings of the tests defining the second factor may have any value on the first factor. However, in these configurations the scores used to define the second factor had loadings in the

range .04 to  $-.19$  on the first factor with a median of  $-.10$  whereas their loadings on the second factor were in the range .55 to .81. It is concluded that the two groups of scores used to define the factors were each homogeneous and that the two groups were orthogonal.

Inspection of the factor loadings and of the table of percentages of persons attempting the last item on each page shows what appears to be a slight systematic effect of practice; but this effect is negligible and could not influence the conclusions regarding the effects of speededness.

TABLE 2  
Factor Loadings

	Form A		Form B		Form C	
	I	II	I	II	I	II
Page 1	[ .69	.01*	[ .60	.08*	[ .73	.05*
Page 2	.58	.03*	.27	.50	.32	.55
Page 3	[ .21	.25	[ .66	.01*	[ .03	.57†
Page 4	[ .68	$-.04^*$	[ .56	.38	[ .66	$-.07^*$
Page 5	.49	.23	.04	.62†	.75	$-.01^*$
Page 6	[ $-.08$	.66†	[ .72	$-.05^*$	[ .40	.37
Page 7	[ .66	$-.05^*$	[ .59	$-.10^*$	[ .62	.08*
Page 8	.57	.04*	.62	.06*	.51	$-.05^*$
Page 9	.60	.13	.54	.34	.46	.37
Page 10	[ $-.10$	.63†	[ $-.05$	.55†	[ $-.10$	.57†
Wrong	$-.51$	.45	$-.46$	.46	$-.55$	.39
Skipped	$-.05$	.04	$-.02$	$-.03$	.13	$-.03$
Attempted	$-.14$	.69†	$-.08$	.72†	$-.13$	.75†
Wrong	$-.52$	.48	$-.58$	.47	$-.58$	.39
Skipped	$-.06$	$-.03$	.06	$-.03$	.16	$-.03$
Attempted	$-.19$	.73†	$-.15$	.81†	$-.17$	.71†
Conduct	.00	.05	.17	.05	.09	.22
Physical Training	.05	$-.04$	.07	.13	$-.01$	$-.02$
Marine Engineering	.42	.14	.36	.14	.35	.28
Mathematics	.36	.11	.26	.08	.24	.15
Electrical Engin.	.32	.06	.21	.00	.28	.14
English	.09	.12	.22	.02	.21	.05
Foreign Languages	.09	.06	.24	.06	.21	$-.07$
Relative Standing	.35	.09	.32	.10	.32	.14
Part Scores on Test						
	(2) .47	.21	(1) .51	.39	(1) .48	.57
	(3) .52	.38	(2) .57	.53	(3) .67	.23
	(5) .58	.38	(5) .55	.49	(5) .50	.55

Note that page numbers indicate the order and not the identity of the pages (which may be determined from Table 1) and that these pages occupy different positions in the various parts of each form. Brackets indicate the extent of the separately timed parts.

\*Scores used in the group defining the first factor.

†Scores used in the group defining the second factor.

The criterion scores for Conduct and Physical Training had low and inconsistent loadings on both factors. The loadings of the other criterion scores on the first factor had a median of .27 and were in the range .09 to .42. These same scores had loadings in the second factor in the range -.07 to .28 with a median at .10. The part-score vector in each of the three analyses which lay closest to the median value for the criterion scores was the vector for the 20-item part. This indicates the superior validity of the moderately speeded test. As was expected, the 30-item parts all had higher loadings on the second factor than did the 20-item parts, and the 10-item parts all had lower loadings than did the 20-item parts.

### *Conclusions*

Two sets of scores on this test form orthogonal clusters. Scores on the first pages of all parts have significant loadings on only the first factor. Scores on the last pages of speeded parts have significant loadings on only the second factor. The criterion scores, grades at the U. S. Naval Academy, are more closely related to scores on first pages than to scores on last pages. The conclusion of this study is that the most valid figure-classification test for predicting grades at the Naval Academy is a moderately speeded test which can be finished by about 70% of the candidates. The score on speeded tests is a function of two orthogonal factors, the factor of ability being more valid than the rate-of-answering factor.

Ordinarily one assumes that those who are able to solve difficult problems in a test will work more quickly than those who cannot. This study suggests that this may not always be the case. The hypothesis is advanced that for some speeded tests not every mark on the answer sheet represents a subject's reasoned conception of an adequate solution. Many of the answers may be guesses, although not necessarily random answers. The group of subjects who reach the end of a speeded test may include some who can solve the problems quickly and also some who answer the problems before they complete the solutions. These subjects may not understand what they are expected to do or they may prefer to guess in order to work quickly. Thus, while it may not always be possible to describe the full effects of speededness by stating the proportion of answers marked, nevertheless the evaluation of a test cannot be complete without some knowledge of the speed at which the subjects were required to work.\*

\*For more complete and detailed information, including tables of intercorrelations and centroid analyses of one of these forms and of an additional form, see Myers, Charles T. The factorial composition and validity of differently speeded tests (second edition) *Research Bulletin* #52-1, Educational Testing Service, Princeton, New Jersey, 1952.

*Manuscript received 5/21/51*

*Revised manuscript received 1/21/52*

## BOOK REVIEW

PAUL S. DWYER. *Linear Computations*. New York: John Wiley and Sons, Inc., 1951. Pp. xi + 344.

*Linear Computations* is addressed to the theoretical and practical considerations of the general problem of finding numerical solutions for systems of linear equations. The author hoped that this volume might serve as a textbook for a course in linear computations and also as a reference book—and from all appearances it seems that both these objectives will be attained. The organization of the material follows sound pedagogical principles which, with the many illustrative problems and the exercises at the close of each chapter, make it especially suitable as a textbook. At the same time, the orderly grouping of methods, the illustrative problems, the annotated references at the end of each chapter, and the detailed index will be appreciated by the research worker who uses it as a reference work in computing techniques.

It would be very presumptuous of the reviewer to attempt to summarize the breadth of the theory and technique covered in *Linear Computations* in a few short pages. Instead, he can but hope that a brief outline of its contents will indicate the scope of the volume.

Introductory remarks, defining and delimiting the problem of linear computations, and general principles of computational design are covered in Chapters 1 and 3. Special emphasis is given in this book (Chapters 2 and 17) to the question of errors in the solutions of simultaneous linear equations when the coefficients themselves are recognized as being subject to error. While some of the classical material on calculations with significant numbers is presented, what is much more important is the introduction of "range numbers" or "approximation-error numbers" for precise results. However, extensive computations with these numbers would be required in determining bounds for the error in the basic linear problems. Therefore, the author recommends the use of "incomplete numbers" with separate estimates of the maximum errors in linear computations such as evaluation of determinants, solutions of simultaneous equations, and determination of elements of the inverse matrix.

An orderly grouping of various methods for solving systems of linear equations is carried out in Chapters 4, 5, and 6, and the relationships among these methods are brought out in Chapter 7. Then applications are made (Chapter 8) to the solution of many sets of simultaneous linear equations derived from the original set by deletion of certain variables or by replacement of the constant terms by new values. The different computational methods are exhibited by schematic arrangement, in general algebraic form, and then followed by numerical illustrations to assist the reader in understanding the development.

Next, an elementary exposition of the theory of determinants is presented (Chapter 9), followed by the application of the various reduction methods to the evaluation of determinants (Chapter 10). Then the reader is introduced to the basic algebra of matrices in Chapter 12. Matrix theory enables the author to develop a more powerful exposition of the solution of simultaneous linear equations, employing the calculation of the inverse matrix (Chapters 13 and 14). Also, certain problems involving the characteristic equation are solved by the method of determinants (Chapter 15). The Hotelling-Bingham-Girshick method is presented for obtaining the inverse of a matrix based on the characteristic equation; and an excellent computational modification of the recently developed recursion formula of J. S. Frame is presented for the determination of the adjoint, the determinant, and the inverse as well as the characteristic equation.

Because of the many diverse approaches to the problem of solving simultaneous

linear equations (and related problems), the author frankly admits that it would be an impossible task to review all of them. He devotes Chapter 16 to the treatment of other methods (e.g., extension and iterative methods) which he does not recommend for use with desk calculators.

While the general problem of finding numerical solutions to systems of linear equations arises in a variety of ways, least squares theory and related statistical studies immediately lead to such mathematical problems. The presentation in this book is of a general mathematical nature rather than in specific statistical terms, although problems in statistics are mentioned. Then in Chapter 18 the author exhibits a number of statistical applications to assist the reader in translating the mathematical results to appropriate statistical results, with emphasis on improvement of the computational design.

Finally, in the concluding chapter, the author looks ahead to procedures for generalizing the techniques presented to non-linear problems; and points to the continued major role that the desk calculator will play (high-speed digital computers notwithstanding) in solving linear problems involving 15 to 20 variables by the elimination methods featured in this book.

The general organization of the text, designation of sections and tables, and over-all typography is such as to enhance the usefulness of this book as a reference work. However, in a volume of this type (with such extensive numerical and algebraic illustrations) it is almost inevitable that errors will appear in the initial printing. A list of such corrections has been prepared by the author and will be distributed by the publishers.

*Linear Computations* is the outgrowth of Professor Dwyer's many years of experience and interest in both the development of complex theory in mathematical statistics and the design of appropriate computational techniques for the application of such theory. The practicability of the book was assured by the author's years of experience in applying the methods to problems in the Statistical Research Laboratory at the University of Michigan, and more recently as a consultant on such matters to the Department of the Army. The use of *Linear Computations* will extend to many fields. It should provide an excellent foundation in numerical analysis—a subject of increasing importance and interest since the rapid development of high-speed electronic digital computers. Psychological statisticians and research psychologists, in particular, will welcome the many computational techniques applicable to their special problems. However, the important implications will not be arrived at by scanning, but diligent work with the material presented can produce very profitable results in psychometrics.

Personnel Research Section,  
Adjutant General's Office

Harry H. Harman

## BOOKS RECEIVED

- FRANZ ALEXANDER AND HELEN ROSS (editors). *Dynamic Psychiatry*. Chicago: University of Chicago Press, 1952, pp. 578.
- EGON BRUNSWIK. *The Conceptual Framework of Psychology*. Chicago: The University of Chicago Press, 1952, pp. iv + 102.
- OSCAR K. BUROS. *Statistical Methodology Reviews, 1941-1950*. New York: John Wiley & Sons, 1951.
- M. F. CLEUGH. *Psychology in the Service of the School*. New York: Philosophical Library, 1951, pp. vii + 183.
- JAMES T. CULBERTSON. *Consciousness and Behaviour*. Dubuque, Iowa: Wm. C. Brown Company, 1950, pp. xii + 210.
- C. A. CURRAN. *Counseling in Catholic Life and Education*. New York: The Macmillan Company, 1952, pp. 462.
- VERNON DAVIES. *Tables Showing Significance of Differences Between Percentages and Between Means*. Pullman: Washington Agricultural Experiment Station Circular No. 151, 1951, pp. 12.
- RICHARD DEWEY AND W. J. HUMBER. *The Development of Human Behavior*. New York: The Macmillan Company, pp. xiv + 762.
- SIEGFRIED FISCHER, M.D. *Principles of General Psychopathology*. New York: Philosophical Library, 1950, pp. xxi + 327.
- J. C. FLUGEL. *A Hundred Years in Psychology*. 2nd ed. London: Duckworth, 1951, pp. 424.
- GEORGE HUMPHREY. *Thinking, An Introduction to Its Experimental Psychology*. New York: John Wiley & Sons, 1951, pp. xi + 331.
- INVITATIONAL CONFERENCE ON TESTING PROBLEMS. Educational Testing Service, Robert Thorndike, Chairman, Princeton, New Jersey, October 28, 1950.
- OSCAR KEMPTHORNE. *The Design and Analysis of Experiments*. New York: John Wiley & Sons, 1952, pp. xi + 631.
- DONALD A. LAIRD AND ELEANOR C. LAIRD. *Sizing Up People*. New York: McGraw-Hill Book Co., Inc., 1951, pp. 270.
- E. F. LINDQUIST. *Educational Measurement*. Washington, D. C.: American Council on Education, 1951, pp. xix + 819.
- HERTA LOEWY. *The Retarded Child*. New York: Philosophical Library, 1951, pp. 160.
- MELVIN H. MARKS. *Psychological Theory*. New York: The Macmillan Company, 1951, pp. xi + 585.
- PAUL MASLOW. *The Individual Through the Rorschach*. Vol. I. New York: Paul Maslow, 1951, pp. 134.

- W. H. MIKESELL, (editor). *Modern Abnormal Psychology*. New York: Philosophical Library, 1950, pp. ix + 880.
- J. C. MOLONEY. *The Battle for Mental Health*. New York: Philosophical Library, 1952, pp. x + 105.
- NIGEL MORLAND. *An Outline of Scientific Criminology*. New York: Philosophical Library, 1950, pp. 287.
- CLARENCE W. NELSON. *Use of Factorial Design in Industrial Relations Research*. Dubuque, Iowa: Wm. C. Brown Company, 1950.
- STANLEY PAYNE. *The Art of Asking Questions*. Princeton, New Jersey: Princeton University Press.
- T. H. PEAR, (editor). *Psychological Factors of Peace and War*. New York: Philosophical Library, 1950, pp. 1-262.
- EDGAR RUBIN. *Experimenta Psychologica*. Copenhagen: Ejnar Munksgaard, 1949, pp. 356.
- JURGEN RUESCH AND GREGORY BATESON. *Communication, the Social Matrix of Psychiatry*. New York: W. W. Norton, 1951, pp. 314.
- D. D. RUNES. *Spinoza Dictionary*. New York: Philosophical Library, 1951, pp. xiv + 309.
- F. V. SMITH. *The Explanation of Human Behavior*. New York: The Macmillan Company, 1952, pp. xi + 276.
- S. S. STEVENS. *Handbook of Experimental Psychology*. New York: John Wiley & Sons, Inc., 1951, pp. 1436.
- JACOB STOLZENBERG, D.D.S. *Psychosomatics and Suggestion Therapy in Dentistry*. New York: Philosophical Library, 1950, pp. xi + 152.
- C. P. STONE (editor). *Annual Review of Psychology*. Stanford: Annual Reviews, 1952, pp. vii + 462.
- G. W. TILTON. *An Educational Psychology of Learning*. New York: The Macmillan Company, 1951.
- C. W. VALENTINE. *Psychology and Its Bearing on Education*. New York: Philosophical Library, 1951, pp. 674.
- D. S. VILLARS. *Statistical Design and Analysis of Experiments for Development Research*. Dubuque, Iowa: Wm. C. Brown Company, 1951, pp. xvii + 455.
- CHARLOTTE WOLFF. *The Hand in Psychological Diagnosis*. New York: Philosophical Library, 1952, pp. xv + 218.

## PSYCHOMETRIC MONOGRAPHS

Three new issues of the Psychometric Monograph Series are in press. Nos. 6, 7, and 8 will be ready for distribution this fall.

The issues of this series are:

Thurstone, L. L. Primary mental abilities.

*Psychometric Monograph No. 1, \$2.00.*

Thurstone, L. L. and Thurstone, Thelma Gwinn. Factorial studies of intelligence.

*Psychometric Monograph No. 2, \$1.50.*

Wolfe, Dael. Factor analysis to 1940.

*Psychometric Monograph No. 3, \$1.25.*

Thurstone, L. L. A factorial study of perception.

*Psychometric Monograph No. 4, \$2.50.*

French, John W. The description of aptitude and achievement tests in terms of rotated factors

*Psychometric Monograph No. 5, \$4.00.*

Degan, James W. Dimensions of functional psychosis.

*Psychometric Monograph No. 6, \$1.50.*

Lord, Frederic. A theory of test scores.

*Psychometric Monograph No. 7, \$2.00.*

Roff, Merrill. A factorial study of tests in the perceptual area.

*Psychometric Monograph No. 8, \$1.50.*

Orders for Psychometric Monographs No. 1 through No. 6 should be sent to The University of Chicago Press, 5750 Ellis Avenue, Chicago 37, Illinois until December 31, 1952. All orders for later issues and all orders for No. 1 through No. 6 beginning January 1, 1953 should be sent to The William Byrd Press, Box 2-w, Richmond 5, Virginia.



